

Signal processing techniques for cardiovascular monitoring applications using conventional and video-based photoplethysmography

THÈSE N° 8493 (2018)

PRÉSENTÉE LE 22 JUIN 2018

À LA FACULTÉ DES SCIENCES ET TECHNIQUES DE L'INGÉNIEUR

GROUPE SCI STI JMV

PROGRAMME DOCTORAL EN GÉNIE ÉLECTRIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Sibylle FALLET

acceptée sur proposition du jury:

Prof. C. Dehollain, présidente du jury

Dr J.-M. Vesin, directeur de thèse

Prof. V. Corino, rapporteuse

Prof. J.-Ph. Couderc, rapporteur

Prof. J.-Ph. Thiran, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2018

*« À l'origine de toute connaissance,
nous rencontrons la curiosité !
Elle est une condition essentielle du progrès. »*

Alexandra David-Néel

Abstract

Photoplethysmography (PPG)-based monitoring devices will probably play a decisive role in healthcare environment of the future, which will be preventive, predictive, personalized and participatory. Indeed, this optical technology presents several practical advantages over gold standard methods based on electrocardiography, because PPG wearable devices can be comfortably used for long-term continuous monitoring during daily life activities. Contactless video-based PPG technique, also known as imaging photoplethysmography (iPPG), has also attracted much attention recently. In that case, the cardiac pulse is remotely measured from the subtle skin color changes resulting from the blood circulation, using a simple video camera. PPG/iPPG have a lot of potential for a wide range of cardiovascular applications. Hence, there is a substantial need for signal processing techniques to explore these applications and to improve the reliability of the PPG/iPPG-based parameters.

A part of the thesis is dedicated to the development of robust processing schemes to estimate heart rate from the PPG/iPPG signals. The proposed approaches were built on adaptive frequency tracking algorithms that were previously developed in our group. These tools, based on adaptive band-pass filters, provide instantaneous frequency estimates of the input signal(s) with a very low time delay, making them suitable for real-time applications. In case of conventional PPG, a prior adaptive noise cancellation step involving the use of accelerometer signals was also necessary to reconstruct clean PPG signals during the regions corrupted by motion artifacts. Regarding iPPG, after comparing different regions of interest on the subject face, we hypothesized that the simultaneous use of different iPPG signal derivation methods (i.e. methods to derive the iPPG time series from the pixel values of the consecutive frames) could be advantageous. Methods to assess signal quality online and to incorporate it into instantaneous frequency estimation were also examined and successfully applied to improve system reliability.

This thesis also explored different innovative applications involving PPG/iPPG signals. The detection of atrial fibrillation was studied. Novel features derived directly from the PPG waveforms, designed to reflect the morphological changes observed during arrhythmic episodes, were proposed and proven to be successful for atrial fibrillation detection. Arrhythmia detection and robust heart rate estimation approaches were combined in another study aimed at reducing the number of false arrhythmia alarms in the intensive care unit by exploiting signals from independent sources, including PPG. Evaluation on a hidden dataset demonstrated that the number of false alarms was drastically reduced while almost no true alarm was suppressed. Finally, other aspects of the iPPG technology were examined, such as the measurement of pulse rate variability indexes from the iPPG signals and the estimation of respiratory rate from the iPPG interbeat intervals.

Keywords : photoplethysmography, heart rate, real-time monitoring, arrhythmias, atrial fibrillation, imaging photoplethysmography, adaptive filter, heart rate variability, ICU alarms, adaptive frequency tracking, organization measures, signal quality, spectral analysis.

Résumé

Les dispositifs médicaux de surveillance basés sur la photopléthysmographie (PPG) joueront probablement un rôle décisif dans l'environnement de santé du futur qui sera préventif, prédictif, personnalisé et participatif. En effet, cette technologie optique présente plusieurs avantages pratiques par rapport aux méthodes standards basées sur l'électrocardiographie, car les dispositifs portables PPG peuvent être utilisés confortablement pour une surveillance continue à long terme pendant les activités de la vie quotidienne. La technique de PPG sans contact basée sur la vidéo, également connue sous le nom de photopléthysmographie par imagerie (iPPG), a aussi attiré beaucoup d'attention récemment. Dans ce cas, le battement cardiaque est mesuré à distance à partir des changements subtils de la couleur de peau résultant de la circulation sanguine, en utilisant une simple caméra. Les technologies PPG/iPPG ont beaucoup de potentiel pour un large éventail d'applications cardiovasculaires. Par conséquent, il existe un besoin important de techniques de traitement du signal pour explorer ces applications et améliorer la fiabilité des paramètres estimés à partir des signaux PPG/iPPG.

Une partie de la thèse est consacrée au développement de techniques de traitement robustes pour estimer la fréquence cardiaque à partir des signaux PPG/iPPG. Les approches proposées ont été conçues à partir d'algorithmes de suivi de fréquence adaptatifs développés précédemment dans notre groupe. Ces outils, basés sur des filtres passe-bande adaptatifs, fournissent des estimations de fréquence instantanée du (des) signal (aux) d'entrée avec un très faible délai, ce qui les rend appropriés pour des applications en temps réel. Pour le PPG classique, une étape d'annulation de bruit adaptative préalable impliquant l'utilisation de signaux d'accéléromètre était également nécessaire pour reconstruire des signaux PPG propres au cours des régions corrompues par des artefacts de mouvement. Concernant l'iPPG, après avoir comparé différentes régions d'intérêt sur le sujet, nous avons supposé que l'utilisation simultanée de différentes méthodes de dérivation des signaux iPPG (c'est-à-dire des méthodes pour dériver les séries temporelles iPPG à partir des valeurs des pixels des images consécutives) serait avantageuse. Des méthodes permettant d'évaluer la qualité du signal en temps réel et de l'intégrer dans l'estimation de la fréquence instantanée ont également été examinées et appliquées avec succès pour améliorer la fiabilité du système.

Différentes applications innovantes impliquant des signaux PPG/iPPG ont également été explorées dans cette thèse. La détection de la fibrillation auriculaire a été étudiée. De nouvelles caractéristiques dérivées directement des formes d'onde PPG, conçues pour refléter les changements morphologiques observés au cours des épisodes arythmiques, ont été proposées et se sont révélées efficaces pour la détection de la fibrillation auriculaire. La détection d'arythmies et les approches robustes d'estimation de la fréquence cardiaque ont été combinées dans une autre étude visant à réduire le nombre de fausses alarmes d'arythmies dans l'unité de soins intensifs en exploitant des signaux provenant de sources indépendantes, y compris des signaux PPG. L'évaluation sur un set de données caché a démontré que le nombre de fausses alarmes a été considérablement réduit alors que presque aucune alarme réelle n'a été supprimée. Enfin, d'autres aspects de la technologie iPPG ont été examinés, tels que la mesure des indices de va-

riabilité cardiaque à partir des signaux iPPG et l'estimation de la fréquence respiratoire à partir des intervalles inter-battement dérivés des signaux iPPG.

Mot clés : photopléthysmographie, fréquence cardiaque, surveillance en temps réel, arythmies, fibrillation auriculaire, photopléthysmographie par imagerie, filtre adaptatif, variabilité de la fréquence cardiaque, alarmes USI, poursuite adaptative de fréquence, mesures d'organisation, qualité du signal, analyse spectrale.

Remerciements

Un travail de thèse ne s'effectue jamais seul, et je tiens à remercier plusieurs personnes pour leur collaboration ou leur soutien durant ce travail.

- Je tiens à remercier tous les membres du jury. Merci à la présidente, Prof. Catherine Dehollain et aux rapporteurs Prof. Valentina Corino, Prof. Jean-Philippe Couderc et Prof. Jean-Philippe Thiran pour leurs commentaires pertinents et les discussions intéressantes lors de l'examen oral.
- La recherche étant toujours un travail d'équipe, je tiens à remercier toute l'équipe du projet Newborn Care, et en particulier Virginie Moser pour notre collaboration dans le cadre de ce projet. Un grand merci aussi à Daniel Ostojic et Tanja Karen pour m'avoir donné l'opportunité d'assister à un enregistrement dans l'unité des soins intensifs néonataux.
- Également un grand merci à Etienne Pruvot et Mathieu Lemay pour le travail effectué dans le contexte du projet miniHolter. Merci Etienne de m'avoir donné l'opportunité d'assister à quelques procédures d'ablation cardiaque et merci à Mathieu pour notre collaboration dans l'analyse des données.
- J'aimerais dire un grand merci à toute la super équipe de l'ASPG et tout particulièrement aux membres avec lesquels j'ai eu la chance de partager un bout de chemin : Andréa Buttu, Adrian Luca, Leila Mirmohamadsadeghi et Sasan Yazdani. Ces années ont été remplies d'un nombre incalculable de bons moments en conférence, au labo et à Sat. Un grand merci à Adrian, Leila et Sasan pour leur précieuse aide dans la rédaction du manuscrit !
- Merci aussi à toute l'équipe du café, Ashkan Yazdani, David Romascano, Damien Matti, Elda Fischì Gomez, Adrian Luca, Leila Mirmohamadsadeghi, Sasan Yazdani et tous les autres pour leur bonne humeur matinale.
- Merci à tous mes amis, qui sauront se reconnaître.
- J'aimerais aussi dire un grand merci à ma famille pour leur soutien et leur encouragements. Ma maman, Agnès Paratte Fallet et son ami Amédé Charmillot et mon frère Matthieu Fallet qui sont à mes côtés, ainsi que mon papa Jean-Michel Fallet qui veille sur moi de là-haut. Merci aussi à mes grand-parents Françoise et Fernand Paratte.
- Un grand merci à mon fiancé, Dominique Rubi, qui a toujours été à mes côtés durant ces quatre ans.
- Et pour finir par le plus important, je tiens à remercier le plus formidable de tous les directeurs de thèse, Jean-Marc Vesin. MERCI pour tout !!! Tes précieux conseils, ton écoute, tes discussions passionnantes, ton humanité et ton humour ont largement contribué à faire de ces quatre années de thèse une magnifique aventure.

Contents

Abstract	v
Résumé	vii
Remerciements	ix
1 Introduction	1
1.1 Motivation and problem statement	1
1.2 Objectives	3
1.3 Organization	4
1.4 Original contributions	5
I Signal Processing Tools for Instantaneous Frequency Estimation	7
2 Instantaneous Frequency Estimation	9
2.1 Motivations	9
2.2 The concept of instantaneous frequency	9
2.3 Adaptive frequency tracking	11
2.3.1 OSC-MSE algorithm	11
2.3.2 SFT algorithm	15
2.3.3 Inclusion of signal quality in adaptive frequency tracking schemes	16
2.4 Spectral purity index	22
2.5 Conclusion	23
II Processing of Photoplethysmographic Signals	25
3 Introduction to Photoplethysmography	27
4 Robust Heart Rate Estimation Based on Adaptive Filtering	31
4.1 Introduction	31
4.1.1 Motivation of the proposed study	32
4.2 Methods	33
4.2.1 Data	33
4.2.2 Adaptive motion artifact reduction	33
4.2.3 Instantaneous heart rate estimation	35
4.2.4 Performance measurement	36
4.3 Results	37
4.3.1 Performance of the presented scheme	37

4.3.2	Effect of adaptive MA reduction	40
4.3.3	Combining inputs	40
4.3.4	Effect of adaptive β and δ	40
4.3.5	Difference between the different types of physical exercises	41
4.3.6	Estimation delay	43
4.3.7	Comparison with the state-of-the-art	43
4.4	Discussion	44
4.5	Conclusion	45
5	Atrial Fibrillation Detection	47
5.1	Introduction	47
5.1.1	Definition of atrial fibrillation	47
5.1.2	Literature review on the use of photoplethysmographic sensors to detect atrial fibrillation	48
5.2	Methods	49
5.2.1	Database	49
5.2.2	Annotation of ECGs	49
5.2.3	Pre-processing	49
5.2.4	Computation of the features	51
5.2.5	Classification	54
5.2.6	Accuracy of the computed IBI series	54
5.3	Results	55
5.3.1	Computed features	55
5.3.2	Classification performance:	55
5.3.3	Comparison with IBI-ECG-derived features	57
5.3.4	Accuracy of the computed PPG-IBI series	57
5.4	Discussion	58
5.5	Conclusion	60
6	A Multimodal Approach to Reduce False Arrhythmia Alarms in the ICU	61
6.1	Introduction	61
6.1.1	Motivations	61
6.1.2	State-of-the-art	61
6.2	Methods	63
6.2.1	Data	63
6.2.2	ECG processing	68
6.2.3	PPG and ABP processing	70
6.2.4	Arrhythmia alarm processing	72
6.2.5	Evaluation metrics	75
6.3	Results	76
6.4	Discussion	77
6.5	Conclusion	79
III	Processing of Video-Derived Photoplethysmographic Signals	81
7	Introduction to Imaging Photoplethysmography	83
7.1	Principle	83
7.2	Review of the iPPG-signal derivation methods	84
7.3	Conclusion	90

8	A Processing Scheme for Real-Time Heart Rate Estimation	93
8.1	Motivations	93
8.2	Database	93
8.3	Methods	95
8.3.1	Tracking and skin segmentation	95
8.3.2	Determination of the best region on the face to derive iPPG signals	96
8.3.3	Evaluation of three algorithms to compute instantaneous HR from the iPPG signal	100
8.3.4	Combining different iPPG signals	106
8.3.5	Signal quality index	110
8.4	Conclusion	115
9	iPPG-Derived Pulse Rate Variability	117
9.1	Motivations	117
9.2	Assessment of iPPG pulse rate variability indexes	117
9.2.1	Introduction	117
9.2.2	Data	118
9.2.3	Methods	118
9.2.4	Results	124
9.2.5	Heartbeat detection accuracy	124
9.2.6	Discussion	130
9.3	Real-time respiratory rate estimation using imaging photoplethysmography inter-beat intervals	131
9.3.1	Introduction	131
9.3.2	Methods	131
9.3.3	Results	132
9.3.4	Discussion	134
9.4	Conclusion	134
10	Conclusion	135
10.1	Summary of achievements	138
10.2	Perspectives	138
	Appendix	141
A	Classifier Evaluation and Ensemble of Classifiers	143
A.1	Evaluation of classifiers	143
A.2	Classifier combination: selection and fusion	145
A.3	The wisdom of crowds	146
B	Video-Based Heart Rate Monitoring in Neonatal Intensive Care Unit: Preliminary Results	149
B.1	Introduction	149
B.2	Methods	150
B.2.1	Database	150
B.2.2	Tracking and skin segmentation	150
B.2.3	iPPG signal derivation and heart rate estimation	151
B.3	Results	152
B.4	Discussion and conclusion	155
	Bibliography	157

Curriculum Vitæ	171
Publications	173

List of Acronyms

AAE	average absolute error
ABP	arterial blood pressure
ACC	accelerometer
AF	atrial fibrillation
ANC	adaptive noise cancelling
AOI	adaptive organization index
ASPG	applied signal processing group
BP	band-pass
bpm	beats per minute
BVP	blood volume pulse
CCD	charge coupled device
DER	detection error rate
DFT	discrete Fourier transform
ECG	electrocardiogram
FA	false alarms
FFT	fast Fourier transform
FSR	fractional spectral radius
HR	heart rate
HRV	heart rate variability
IBI	interbeat intervals
ICA	independent component analysis
ICU	intensive care unit
iPPG	imaging photoplethysmography
IR	infrared
LED	light-emitting diode
LMS	least mean squares

LOA	limits of agreements
MA	motion artifacts
NFB	notch filter bank
NIR	near-infrared
NLMS	normalized least mean squares
PCA	principal component analysis
PE	permutation entropy
PPG	photoplethysmography
PRV	pulse rate variability
PSD	power spectral density
RMSE	root-mean-square error
ROI	region of interest
SE	spectral entropy
SNR	signal-to-noise ratio
SPI	spectral purity index
SR	sinus rhythm
SVD	singular value decomposition
SWASVD	sliding-window adaptive singular value decomposition
VA	ventricular arrhythmias

List of Tables

4.1	Database description. BP: blood pressure.	33
4.2	Error-1 results for the four different configurations (a-d) of the presented HR estimation scheme.	38
4.3	Error-2 results for the four different configurations (a-d) of the presented HR estimation scheme.	39
4.4	Number of times each channel combination led to best heart rate estimation. . . .	42
4.5	Comparison with existing studies (error-1).	43
5.1	Number of epochs of SR, AF and VA, for each patient.	50
5.2	AUC values for the different features.	56
5.3	Classification performance AF against SR.	56
5.4	Classification performance AF against VA.	57
5.5	Classification performance AF against (SR&VA).	57
5.6	Classification performance for ECG	58
5.7	DER of the detected heartbeats in PPG signal.	58
6.1	Definition of the five alarm types [90, 91].	63
6.2	Algorithms and selected parameters for HR estimation using PPG/ABP.	72
6.3	Results obtained on the hidden-test dataset.	76
6.4	Results for the training set.	76
6.5	Comparison between our scheme and the voting algorithm.	77
6.6	Percentage of processed waveforms for each type of alarm in the training set. . .	77
6.7	Linear discriminant analysis for ventricular tachycardia alarms.	79
7.1	Summary of the state-of-the-art iPPG signal derivation methods.	91
8.1	iPPG database acquisition	95
8.2	Average percentage of power in the HR band for the different regions.	99
8.3	AAE and LOA results for the three algorithms and for the sequences recorded in visible light.	105
8.4	AAE and LOA results for the three algorithms and for the sequences recorded in the dark.	105
8.5	Number of times each iPPG-input resulted in the best/worst HR estimation. . . .	108
8.6	Performance comparison between the original scheme and the SQI-based scheme for the real-time configuration.	113
8.7	Performance comparison between the original scheme and the SQI-based scheme for the aligned configuration.	113
9.1	Performance of heartbeat detection on the PPG signal, for each sequence.	124

9.2	Mean heartbeat detection performance metrics for all sequences and the different iPPG signals.	125
9.3	Performance of heartbeat detection on the iPPG _{SSR} signal, for each sequence. . .	125
9.4	Summary of the agreement between the HRV and PPG-PRV indexes for <i>respiration</i> sequences. * indicates that the p-value was lower than 0.01.	126
9.5	Summary of the agreement between the HRV and PPG-PRV indexes for <i>handgrip</i> sequences.	126
9.6	Summary of the agreement between the HRV and iPPG-PRV indexes for <i>respiration</i> sequences.	127
9.7	Summary of the agreement between the HRV and iPPG-PRV indexes for <i>handgrip</i> sequences.	127
9.8	Summary of the agreement between the PPG-PRV and iPPG-PRV indexes for <i>respiration</i> sequences.	128
9.9	Summary of the agreement between the PPG-PRV and iPPG-PRV indexes for <i>handgrip</i> sequences.	128
9.10	The mean correlations between the IBI of the three iPPG channels and the ECG inter-beat intervals over all subjects.	132
9.11	The mean (\pm standard deviation) error in breaths-per-minute of the ECG and iPPG _{green} estimates compared to the reference over all subjects.	134
A.1	Confusion matrix of hypothetical model trained for a binary classification problem.	144
B.1	The neonate database subset.	151
B.2	Results.	153

List of Figures

1.1	Number of publications and citations for PPG studies (1994-2014).	2
2.1	General configuration of the adaptive BP filter proposed in [19].	11
2.2	Amplitude and phase response of the BP filter used in the OSC-MSE algorithm. .	13
2.3	Frequency estimation using the OSC-MSE algorithm ($\delta = 0.8$) and corresponding RMSE.	13
2.4	The estimation delay of the OSC-MSE algorithm as a function of δ	14
2.5	Configuration of the multi-signal adaptive frequency tracker (OSC-MSE-W). . .	15
2.6	Amplitude and phase response of the BP filter used in the SFT algorithm. . . .	17
2.7	Bias, variance and AAE for the OSC-MSE-W and OSC-MSE-W-SQI algorithms. .	18
2.8	Example of two synthetic input signals.	19
2.9	Adaptive parameters as functions of the average SPI.	20
2.10	Bias, variance and AAE for the STF and SFT-SPI algorithms.	20
2.11	Bias, variance and AAE for the STF-W and SFT-W-SPI algorithms, with simulated bad-quality segments.	21
2.12	Frequency estimation using the SFT-W-SPI algorithm on a synthetic signal. . . .	21
2.13	The SPI of a synthetic signal with three different regimes.	23
3.1	The two possible operational configurations of a PPG sensor system.	27
3.2	Variation in light attenuation by tissue.	28
3.3	PPG waveform and pulse landmarks frequently used.	29
3.4	Example of waveforms acquired in two healthy subjects at different body locations. .	29
4.1	Example to illustrate how motion artifacts can corrupt the PPG waveforms. . . .	32
4.2	General configuration of adaptive filters.	34
4.3	Concept of adaptive noise cancelling.	34
4.4	Example of SPI for a PPG signal (dataset-1, record 02). (a) PPG signal; (b) SPI. .	36
4.5	Flow diagram of the proposed HR estimation scheme.	36
4.6	Bland-Altman plots for dataset-1 (a) and dataset-2 (b).	40
4.7	Effect of adaptive MA reduction on a data segment (record 02, dataset-1). . . .	41
4.8	HR estimation with and without adaptive MA reduction (record-12).	41
4.9	Relationships between the error-1 and δ/β for three records.	42
4.10	Error-1 results for the different types of physical activities.	42
5.1	Example of ECG showing AF. Figure from [55].	47
5.2	Removal of PPG segments contaminated by motion artifacts.	51
5.3	Boxplot of the obtained AOI values for the four classes.	52
5.4	SE values obtained for a sinusoid contaminated with different level of white Gaussian noise.	53

5.5	Computation of the AOI feature for the different types of rhythms: SR, AF and VA (from the left to the right).	55
5.6	Example of PPG waveforms (a,c), their PSD (b,d) and the resulting SE values. Top row: SR; bottom row: AF.	56
5.7	Boxplots for all features, computed from all the labeled PPG epochs.	57
5.8	Example of a VA episode. (a) ECG (lead II); (b) PPG signal.	60
6.1	The general framework to determine the validity of an alarm.	63
6.2	Example of a true asystole alarm (tape a142s).	64
6.3	Example of a false asystole alarm (tape a134s).	64
6.4	Example of a true bradycardia alarm (tape b455l).	65
6.5	Example of a false bradycardia alarm (tape b332s).	65
6.6	Example of a true tachycardia alarm (tape t174s).	66
6.7	Example of a false tachycardia alarm (tape t409l).	66
6.8	Example of a true ventricular tachycardia alarm (tape v648s).	67
6.9	Example of a false ventricular tachycardia alarm (tape v169l).	67
6.10	Example of a true ventricular flutter/fibrillation alarm (tape f544s).	68
6.11	Example of a false ventricular flutter/fibrillation alarm (tape f121l).	68
6.12	The block diagram of AMM.	69
6.13	AMM performance on a tape with bigeminy (v421l).	69
6.14	AMM performance on an ECG (tape a527l) from the training set.	70
6.15	Example of SPI for a true ventricular tachycardia (record v194s).	71
6.16	Interest of using multiple inputs for adaptive frequency tracking.	72
6.17	An example of HR estimation using the OSC-MSE-W algorithm during an extreme tachycardia episode. (a) PPG waveform; (b) Instantaneous HR.	73
6.18	Asystole alarm processing.	74
6.19	Extreme bradycardia alarm processing.	74
6.20	Extreme tachycardia alarm processing.	74
6.21	Ventricular flutter/fibrillation alarm processing.	75
6.22	Ventricular tachycardia alarm processing.	75
6.23	Two examples of PPG signal behavior during true ventricular tachycardia episodes.	78
7.1	Skin reflection model. Image from [96], © 2016 IEEE.	83
7.2	Basic principle of iPPG signal derivation, as introduced in [5].	84
7.3	Principle of BSS-based iPPG signal derivation methods.	85
7.4	The different steps involved in the POS method [96].	87
7.5	The SSR method described in [106].	89
8.1	iPPG acquisition protocol: respiration modulation. brpm: breaths per minute.	94
8.2	iPPG acquisition protocol: handgrip isometric exercise.	94
8.3	The experimental setup used for the acquisition of the <i>adult database</i>	95
8.4	A sample ROI tracked on the subject forehead and the skin segmentation mask.	96
8.5	A sample ROI encompassing the whole face and the skin segmentation mask.	96
8.6	Division of the face area into 260 small ROIs.	97
8.7	Percentages of power at true HR for the different face regions and channels.	98
8.8	Normalized color maps averaged over all subjects.	98
8.9	Time evolution of the power at HR frequency in the different zones.	99
8.10	SWASVD algorithm. Image from [128], © 2004 IEEE.	102
8.11	SWASVD-based scheme for HR estimation.	102
8.12	The different steps of the NFB algorithm.	103
8.13	Sample of iPPG _{NIR} signal derived from a sequence acquired in total darkness.	104

8.14	Estimated HR values and ground-truth for a recording of the <i>handgrip</i> subset.	104
8.15	AAE for HR estimation from iPPG for all sequences.	105
8.16	Sample of the iPPG signals derived with the different methods for a recording from the <i>respiration</i> subset. ECG is also displayed as visual reference.	107
8.17	Sample of the iPPG signals derived with the different methods for a recording from the <i>motion</i> subset. ECG is also displayed as visual reference.	108
8.18	Individual AAE results for each sequence and each input signal, for the <i>static</i> subset. The estimation delay was compensated.	109
8.19	Individual AAE results for each sequence and each input signal, for the <i>motion</i> subset. The estimation delay was compensated.	109
8.20	The time evolution of the DI for a sequence taken in the <i>motion</i> subset.	111
8.21	Computation of the proposed iPPG SQI.	112
8.22	The SQI for a sequence from the <i>motion</i> subset (forehead ROI).	112
8.23	Individual AAE values and percentages of data considered as not reliable for HR estimation for the <i>motion</i> subset.	114
8.24	HR estimated with and without using the SQI for a sequence of the <i>motion</i> subset.	114
9.1	ECG and pulse-wave signal.	119
9.2	Exemple to illustrate PPG baseline removal and heartbeat detection.	120
9.3	Example of a PPG segment corrupted by motion artifacts.	120
9.4	Example to illustrate heartbeat detection on the iPPG signal.	121
9.5	The three IBI series for a subject (sequence 1, <i>respiration</i> protocol).	122
9.6	The three IBI series for a subject (sequence 6, <i>handgrip</i> protocol).	122
9.7	Agreement between ECG-HRV and PPG-PRV time domain measures.	129
9.8	Agreement between ECG-HRV and iPPG-PRV time domain measures.	129
9.9	The iPPG _{green} (a) quality index and (b) respiratory rate, ECG-derived respiratory rate and reference for subject 1.	133
9.10	The iPPG _{green} (a) quality index and (b) respiratory rate, ECG-derived respiratory rate and reference for subject 2.	133
A.1	An example classifier selection technique illustrated in a two attribute subspace.	146
A.2	Typical framework of classifier fusion.	147
B.1	Experimental setup used for the database acquisition at the USZ NICU.	150
B.2	Example of iPPG signals and their PSD for an 8-second data segment.	152
B.3	Sequence P3-R1. Estimated and reference HRs (top) and <i>DI</i> measure.	153
B.4	Sequence P6-R3. Estimated and reference HRs (top) and <i>DI</i> measure.	154
B.5	Sequence P13-R1. Estimated and reference HRs (top) and <i>DI</i> measure.	154

Introduction

1.1 Motivation and problem statement

Cardiovascular monitoring techniques have considerably evolved in the last decades, resulting in an increased surveillance capability for high-risk subjects [1]. One of the main reasons of this sudden evolution is the appearance of wearable biosensors, which enable continuous monitoring in the hospital and at home. The development of such devices was made possible by recent advances in sensing technologies, embedded systems, wireless communication technologies and miniaturization. In addition to enable monitoring during daily life activities, wearable biosensors also facilitate it at the hospital by reducing the amount of cables, which increase the risk of tripping and falling [2].

The gold standard method for non-invasive cardiovascular monitoring is electrocardiography. In electrocardiography, electrodes placed on the patient's limbs and on the surface of the chest are used to measure the electrical potentials generated by the heart electrical activity. The resulting signal, named electrocardiogram (ECG), can be used for numerous applications, including heart rate estimation, heart rate variability analysis, arrhythmia detection and breathing rate estimation. ECG has high diagnostic capabilities, but is not the best candidate for continuous monitoring during daily life activities. Indeed, ECG requires the placement of multiple electrodes on the chest and the cables are encumbering. In addition, adhesive electrodes can lead to some issues such as skin infection and signal deterioration over long periods of time. For this reason, ECG-based long term monitoring should be avoided [3]. There is a growing interest for alternative ECG methods based on non-contact electrodes that can be integrated into smart textiles for instance.

Instead of measuring the electrical activity of the heart, one can retrieve information about the cardiovascular system from the blood volume changes in the microvascular bed of tissue, detected with an optical sensor. This technique, known as photoplethysmography, has widespread clinical applications (see Chapter 3 for a thorough description). This dissertation focuses on the processing of photoplethysmographic (PPG) signals in the context of different monitoring applications. Compared to ECG, PPG technology presents several advantages related to its usage. Indeed, a PPG system requires only a few opto-electronic components (an illumination source and a photodetector) and can easily be embedded in comfortable wearable devices such as wrist-type, ear-worn or finger ring sensors. Moreover, while ECG signals have to be measured from several body sites, PPG signals can be retrieved from a single body site. In addition, PPG sensors are usually low-cost. For all these reasons, PPG devices have become very popular in the last two decades. This sudden surge of interest in the PPG technology is demonstrated by the considerable increase in the number of publications and citations for PPG studies, as shown in

Figure 1.1.

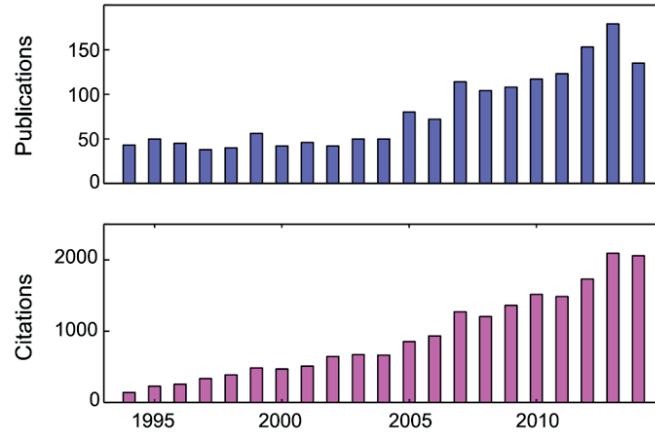


Figure 1.1: Number of publications and citations for PPG studies (1994-2014). Image from [4], © 2016 IEEE.

Contactless video-based PPG techniques have also attracted much attention recently. In 2008, it was shown for the first time that the cardiac pulse can be remotely measured from the subtle skin color changes, induced by the blood circulation, using a simple video camera [5]. This completely contactless and low cost method, referred to as imaging photoplethysmography (iPPG) in this dissertation, opens new horizons for home monitoring and telemedicine.

Numerous challenges emerge from these new PPG-based cardiovascular monitoring techniques. From a physiological point of view, it is important to fully understand the relationship between the studied pulse signal (PPG or iPPG) and the ECG waveform, and to what extent cardiovascular parameters usually monitored with ECG can be monitored with PPG. In this regard, many studies recently reported the accuracy with which some cardiovascular parameters, such as heart rate, heart rate variability indexes and respiration rate could be derived from PPG signals. The detection of arrhythmias from the PPG signal is also a topic of growing interest. Indeed, PPG-based monitoring systems are more user-friendly than conventional ECG-based systems. Firstly, PPG enables long-term monitoring, which is necessary for the diagnostic of some cardiac dysrhythmias. Secondly, PPG-based systems are affordable and comfortable to wear, and could therefore be used to screen large population for preventive purposes. Finally, arrhythmia detection from PPG could also be used as a complementary information in the intensive care unit, where the high false alarm rates are of serious concern.

PPG-based cardiovascular monitoring is a challenging area from a signal processing point of view for many reasons. One of the crucial issues is the quality of the data acquired with such wearable/contactless PPG systems. Indeed, PPG signals are very sensitive to patient and/or probe-tissue movement artifacts [6]. Algorithms are therefore needed to deal with these bad-quality episodes. For instance, heart rate estimation during physical exercise is not a trivial task and methods based on heartbeat detection in the time domain usually fail during these perturbed periods. Fourier-based methods are more appropriate to perform instantaneous frequency estimation in such cases, but they entail complex peak selection process and induce an estimation delay of half the length of the window used to compute the spectrum. Despite the large number of publications on this topic, new methods to reconstruct clean signals from corrupted data as well as approaches to improve the reliability of instantaneous frequency estimation are still under investigation.

Video-derived PPG signals are equally, if not more, concerned by this problematic. Indeed, signal stability is usually low for this contactless method. Therefore, both image and signal processing algorithms need to be robust against numerous disturbances such as changes in ambient illumination and movements.

Banaee *et al.* have addressed the topic of data mining tasks for wearable sensors in a literature review [7]. These tasks were divided into three categories: 1) anomaly detection, 2) prediction and 3) diagnosis/decision making. Anomaly detection consists in detecting unusual patterns. This task has to be performed in real-time in online health monitoring systems so that alarms are triggered as soon as vital signs become abnormal. Prediction involves the use of a predictive model to prevent chronic problems by identifying events which have not yet occurred. Diagnosis/decision making is linked to anomaly detection, as abnormal patterns usually provide valuable information about diagnosis. The following challenges related to the different tasks were brought out:

- Simultaneous exploitation of multiple measurements of vital signs.
- Increase in the level of trust of the system used for the decision making process.
- Development of real-time health systems.

These are exactly the challenges that were dealt with in the algorithmic developments of the present dissertation. Indeed, the combination of multiple waveforms/parameters often increases the reliability of a given system, provided that the underlying combination process is appropriate. System's level of trust is the major concern when dealing with PPG/iPPG signals, given that they are easily deteriorated by motion artifacts. Finally, many challenges are raised by the design of real-time approaches since they require the use of causal processing techniques with low computational complexity.

To summarize, PPG-based monitoring systems are promising tools that will probably play a determinant role in the future healthcare system, which is usually described as preventive, predictive, personalized, pervasive, participatory and patient-centered [3]. The sudden increase in popularity of PPG-based monitoring devices is accompanied by a substantial need for novel signal processing techniques to achieve reliable vital sign estimation, as well as to investigate their potential for a wide range of applications.

1.2 Objectives

The present dissertation aims at developing signal processing approaches for different cardiovascular monitoring applications based on conventional PPG or video-based PPG. The applications studied can be divided into two categories: 1) continuous estimation of vital signs and 2) arrhythmia detection. Regarding the development of algorithms for vital signs monitoring, the aim is to meet an accuracy comparable to the one obtained with the reference sensor (i.e. ECG). In addition, the approaches proposed are developed with a special emphasis on real-time/online processing and low computational complexity, to facilitate a potential embedding into device processors.

Regarding the processing of conventional PPG signals, the objectives are:

- To develop robust heart rate estimation schemes suitable to monitor the subjects during their daily life activities:
 - To reconstruct clean PPG signals during periods corrupted by motion artifacts using adaptive noise reduction techniques.

- To evaluate the potential benefits of combining different waveforms in adaptive frequency tracking schemes.
- To assess the potential of PPG-derived features in the detection of cardiac arrhythmias:
 - To develop new metrics to quantify the organization/complexity of PPG signals captured during atrial fibrillation.
 - To detect life-threatening arrhythmias in the intensive care unit and suppress false alarms.

Regarding the processing of iPPG signals, the objectives are:

- To investigate different aspects of this innovative technology. More specifically:
 - To compare the suitability of different regions of interest on the face to derive iPPG signals.
 - To evaluate the potential benefits of combining different existing iPPG-signal derivation methods in a real-time heart rate estimation scheme.
 - To enhance the reliability of heart rate estimation by using a signal quality index.
 - To assess the reliability of iPPG-derived heart rate variability indexes during non stationary conditions.

1.3 Organization

The dissertation is divided into three parts. The first part describes adaptive frequency tracking algorithms that were used to estimate heart rate from PPG/iPPG signals. The second part focuses on the processing of conventional PPG signals for various applications. The third part is dedicated to the development of processing schemes and analysis techniques for video-based vital sign monitoring applications.

Part I: Signal processing tools for instantaneous frequency estimation: This first part presents the tools that were used for instantaneous frequency estimation throughout this dissertation. The concept of instantaneous frequency and its importance in physiological signals is first discussed. Then, two adaptive frequency algorithms are introduced and discussed. These algorithms, which can combine multiple input signals, enable the estimation of the instantaneous frequency with a very small delay. Two extensions of these algorithms are then proposed in order to increase the robustness of frequency estimation by taking signal quality into account. The spectral purity index, which was used in two applications presented in this dissertation, is also introduced.

Part II: Processing of photoplethysmographic signals: Different applications involving particular processing schemes of PPG signals are studied in this part. Chapter 3 gives a short introduction to PPG. Chapter 4 is devoted to the development of a processing scheme to estimate heart rate from PPG signals obtained from subjects performing physical exercise. Chapter 5 is dedicated to atrial fibrillation detection from PPG signals acquired during cardiac ablation procedures. Chapter 6 combines the aspects studied within the two preceding chapters and proposes a multi-modal approach to reduce the number of false arrhythmia alarms in the intensive care unit.

Part III: Processing of video-derived photoplethysmographic signals: This part is dedicated to the processing of video sequences for the remote measurement of cardiovascular parameters. Chapter 7 introduces the iPPG technology and describes various methods that can be used to derive iPPG signals from video-sequences. Chapter 8 describes the development of a complete processing scheme for heart rate estimation from video-sequences. Several parameters such as the selection of the face region of interest to derive the iPPG signals, the iPPG signal derivation

techniques and the waveform quality were taken into consideration for the development of this processing scheme. Chapter 9 is devoted to the measurement of pulse rate variability indexes from the iPPG signals and the estimation of respiratory rate from the iPPG interbeat intervals.

1.4 Original contributions

The main contributions of this work are:

Part I: Signal processing tools for instantaneous frequency estimation:

- Development of two mechanisms to include the information about signal quality into the adaptive frequency tracking algorithms.

Part II: Processing of photoplethysmographic signals:

- Design of a two-step heart rate estimation scheme for subjects performing physical exercise, consisting of 1) an adaptive noise cancellation step and 2) an adaptive frequency tracking step.
- Development of a multi-modal approach based of innovative ECG-, PPG-, and arterial blood pressure (ABP)-derived measures and an ensemble of logic decision rules to decrease the number of false arrhythmia alarms in the intensive care unit.
- Introduction of novel PPG-wave measures reflecting signal organization/complexity and assessment of their potential to discriminate between sinus rhythm, atrial fibrillation and ventricular arrhythmias.

Part III: Processing of video-derived photoplethysmographic signals:

- Assessment of the spatial distribution of heart rate related information on the face.
- Development of a real-time heart rate estimation scheme taking advantage of multiple existing iPPG derivation methods and combining them in an adaptive frequency tracking scheme.
- Development of an iPPG signal quality index computed from image and signal features that is suitable for real-time applications.
- Assessment of selected iPPG-derived pulse rate variability parameters in non-stationary conditions and comparison of these parameters with the same parameters derived from the ECG.

Part I

Signal Processing Tools for Instantaneous Frequency Estimation

Instantaneous Frequency Estimation

2

The present chapter aims at introducing the frequency estimation tools that were used in various applications in the framework of this thesis. The concept of instantaneous frequency is described in Section 2.2. Two adaptive frequency tracking algorithms are then introduced in Sections 2.3.1 and 2.3.2. Their multi-signal extensions are also presented, as well as developed extensions aiming at taking the signal quality into account in frequency estimation process. The concept of spectral purity index is introduced in Section 2.4.

2.1 Motivations

Many physiological phenomena are driven by specific rhythms. These rhythms cover a large range of frequencies, with periods ranging from a few milliseconds up to several months [8]. The study of these physiological rhythms provides valuable information about body functions and their underlying regulation mechanisms. Moreover, the interactions between the different systems can also be investigated through the relationship between the different rhythms, such as synchronization and mutual entrainment. Beyond the understanding of these fundamental aspects, particular frequencies can be associated to specific physical states for monitoring purposes (such as breathing frequency, heart rate, sleep stages, etc.). From a signal processing point of view, the extraction and characterization of a number of these rhythms remain a challenge for various reasons. Firstly, most of the physiological processes are non-stationary, making all the processing tools developed for the stationary case not suitable. Secondly, several oscillations can be present simultaneously. In addition, bio-signals recorded in a non-invasive manner from wearable devices can be easily polluted by artifacts. Finally, when the aim is to monitor the patient, the frequency estimation has to be performed with the smallest possible time-delay, so that an alarm can be triggered if needed. For all the aforementioned reasons, researchers continue to investigate new tools to improve the reliability of frequency estimation in such challenging conditions.

2.2 The concept of instantaneous frequency

Frequency is defined as the number of oscillations observed per unit of time. However, many processes are nonstationary, i.e. processes whose mean, correlation function, and higher-order moments are time-varying [9], and therefore the spectral characteristics vary over time. In that case, the concept of instantaneous frequency is useful, as it reflects the frequency at a given time.

In order to define the instantaneous frequency, an introduction to the concept of the analytic signal is needed. The analytic signal $x_a(t)$ is a complex signal defined as:

$$x_a(t) = x(t) + ix_h(t) \quad (2.1)$$

with $x_h(t)$, the Hilbert transform of $x(t)$. The Hilbert transform is a linear operator, which can be expressed as the convolution of the input signal $x(t)$ and the impulse response of the quadrature filter. The quadrature filter can be seen as an all pass filter providing a phase shift of $-\pi/2$ radians for the positive frequency components and $\pi/2$ radians for the negative frequency components. Its frequency response is:

$$H(e^{j\omega}) = \begin{cases} j, & -\pi < \omega < 0 \\ 0, & \omega = 0 \\ -j, & 0 < \omega < \pi \end{cases} \quad (2.2)$$

and the corresponding impulse response is:

$$h(t) = \frac{1}{\pi t} \quad (2.3)$$

The Hilbert transform is the response of the quadrature filter to a real input signal $x(t)$:

$$x_h(t) = x(t) * h(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (2.4)$$

When dealing with discrete-time systems, the impulse response of equation 2.3 becomes:

$$h[n] = \begin{cases} \frac{2}{\pi n}, & \text{for odd } n \\ 0, & \text{for even } n \end{cases} \quad (2.5)$$

In practical situations, the Hilbert filter can be approximated with a finite impulse response (FIR)¹.

In order to introduce the notions of instantaneous amplitude and frequency, the analytic signal can be expressed in polar coordinates as follows:

$$x_a(t) = A(t)e^{j\phi(t)} \quad (2.6)$$

with $A(t) = |x_a(t)|$, the instantaneous amplitude and $\phi(t) = \arg(x_a(t))$, the instantaneous phase. The instantaneous angular frequency is defined as the time derivative of the unwrapped phase:

$$\omega(t) = \frac{d\phi(t)}{dt} \quad (2.7)$$

Even though the instantaneous frequency can be obtained by differentiating the instantaneous phase, this approach requires the oscillation to be narrow-band to lead to meaningful results. This method is therefore not well suited for biomedical signals, which are often characterized by more than one frequency component.

1. For more detailed explanations of the Hilbert transform and design of Hilbert Transformers, see [10]

2.3 Adaptive frequency tracking

Throughout this manuscript, the term adaptive frequency tracking will refer to methods based on adaptive band-pass (BP) filters designed to track time-varying oscillations contained in the input signal.

FIR adaptive line enhancers have been previously used in adaptive frequency tracking [11–13]. Although good performances were achieved, observed convergence rates were rather slow. On the other hand, adaptive infinite impulse response (IIR) BP filters have a faster convergence rate. Different schemes based on adaptive IIR filters have been proposed to perform adaptive frequency tracking, with two main types of adaptive mechanisms: non-gradient updating and mean square error (MSE) [14–18]. A new type of coefficient updating mechanism based on the discrete oscillator model was proposed by Liao [19]. This coefficient updating mechanism has the advantage of having a low computational complexity and being independent of the structure chosen for the line-enhancement filter. Figure 2.1 shows the general configuration of an adaptive IIR BP filter.

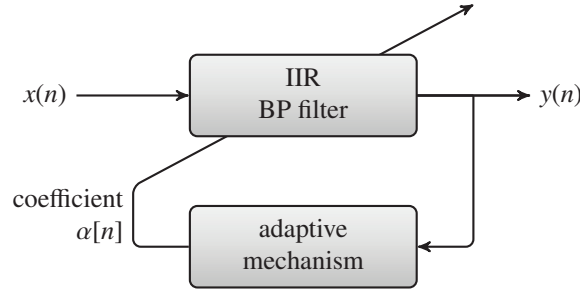


Figure 2.1: General configuration of the adaptive BP filter proposed in [19].

The next two subsections (Sections 2.3.1 and 2.3.2) focus on two adaptive frequency tracking algorithms. These algorithms, derived from the algorithms proposed by Liao [19], have been extended by former PhD researchers of our group, in particular to work in the multi-signal case^{2 3} [20–23]. The multi-signal versions of these algorithms take several oscillatory input signals and track their common instantaneous frequency component. In the biomedical field, in which it is quite frequent to obtain redundant information across the sensors, these tools have revealed to be very useful. Moreover, the signal combination often results in a more robust estimate than when a single input is provided.

2.3.1 OSC-MSE algorithm

The discrete oscillator-based adaptive BP filter was proposed by Liao [19]. An input signal of the following form is considered:

$$x[n] = s[n] + b[n] \quad (2.8)$$

with $s[n]$, a sinusoid of the form $s[n] = A_0 \sin(\omega_0 n + \phi_0)$ and $b[n]$, an additive independent and identically distributed noise. The transfer function of the time-varying BP filter is expressed as follows:

$$H(z; n) = \frac{1 - \beta}{2} \frac{1 - z^{-2}}{1 - \alpha[n](1 + \beta)z^{-1} + \beta z^{-2}} \quad (2.9)$$

2. Y. Prudat, J.-M. Vesin (Dir.). *Adaptive frequency tracking and application to biomedical signals*. PhD Thesis EPFL, n° 4447 (2009).

3. J. Van Zaen, J.-M. Vesin (Dir.). *Efficient Schemes for Adaptive Frequency Tracking and their Relevance for EEG and ECG*. PhD Thesis EPFL, n° 5476 (2012).

with $\alpha[n]$, controlling the central frequency of the BP filter and β ($0 \ll \beta < 1$), a factor related to the filter bandwidth. The adaptive mechanism used to update the central frequency of the BP filter of equation 2.9 at each time step requires the minimization of an MSE term, which is derived from the real discrete oscillator equation described as:

$$s[n] = 2 \cos(\omega_0)s[n-1] - s[n-2] = 2\alpha_0s[n-1] - s[n-2] \quad (2.10)$$

with $\alpha_0 = \cos(\omega_0)$. As illustrated in Figure 2.1, the filter output, $y[n]$, is used in the adaptive mechanism. $y[n]$ is defined by the following difference equation:

$$y[n] = (1 + \beta)\alpha[n]y[n-1] - \beta y[n-2] + \frac{1 - \beta}{2}(x[n] - x[n-2]) \quad (2.11)$$

The following cost function can be derived from equation 2.10:

$$J[n] = E\{(y[n] - 2\alpha[n+1]y[n-1] + y[n-2])^2\} \quad (2.12)$$

The optimal value of the adaptive parameter α can be found by setting to zero the derivative of this cost function, with respect to $\alpha[n+1]$. This leads to the following expression for $\alpha[n+1]$:

$$\alpha[n+1] = \frac{E\{y[n-1](y[n] + y[n-2])\}}{2E\{(y[n-1])^2\}} \quad (2.13)$$

which can be recursively estimated in practice using an exponentially weighted time-average:

$$\alpha[n+1] = \frac{Q[n]}{2P[n]} \quad (2.14)$$

with

$$\begin{aligned} Q[n] &= \delta Q[n-1] + (1 - \delta)y[n-1](y[n] + y[n-2]) \\ P[n] &= \delta P[n-1] + (1 - \delta)(y[n-1])^2 \end{aligned} \quad (2.15)$$

where the convergence rate can be adjusted with a forgetting factor δ . Finally, the instantaneous frequency $\omega[n+1]$ is defined as:

$$\omega[n+1] = \arccos(\alpha[n+1]) \quad (2.16)$$

Figure 2.2 shows the amplitude and phase responses of the BP filter described by the transfer function of equation 2.9, for different values of the β parameter and for $\alpha = \cos(0.5\pi)$. A value of β closer to one entails a narrower filter.

As an example to illustrate the convergence speed as a function of δ , the algorithm was applied to a 1800-sample long sinusoid with additional Gaussian white noise (SNR = 20 dB). At sample $n=600$, the frequency of the sinusoid changes from 0.1π to 0.2π . For 50 values of the parameter δ equally spaced between 0.8 and 0.995, the time (i.e. number of samples) to converge to the second frequency value was estimated. For this purpose, the root-mean-square error (RMSE) of the estimated frequency was computed on a 10-sample sliding window. The estimation delay was defined as the first point after the frequency transition for which the RMSE was lower than a fixed threshold of 0.0001, as illustrated in Figure 2.3. Results were averaged over 1000 runs for each value of δ . Figure 2.4 shows the resulting estimation delay, in samples, for the different values of δ . It can be seen that the delay increases exponentially with δ .

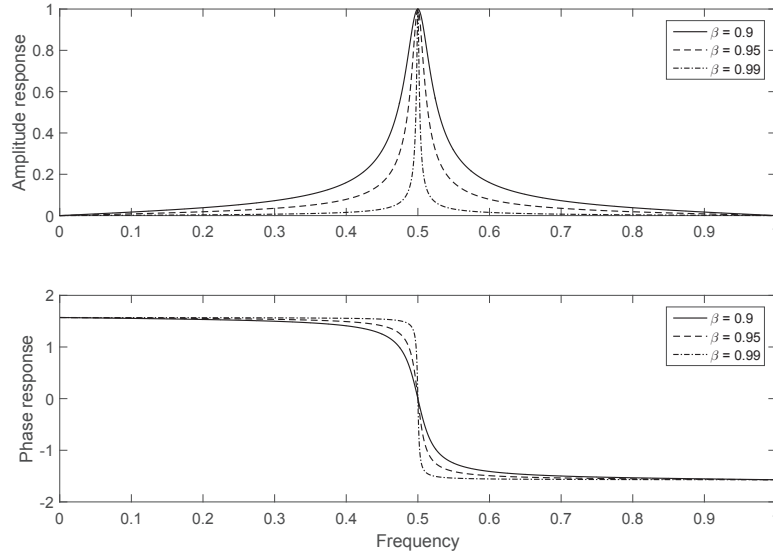


Figure 2.2: Amplitude and phase response of the BP filter used in the OSC-MSE algorithm (see equation 2.9), for different values of β .

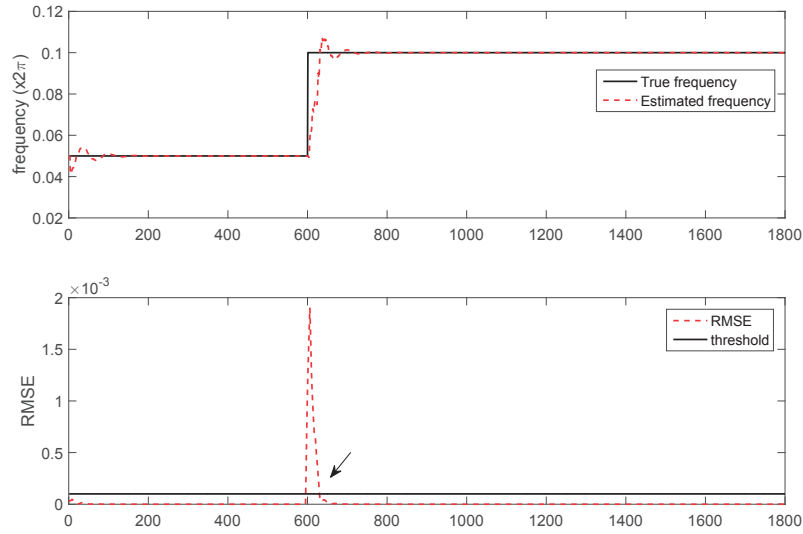


Figure 2.3: Frequency estimation using the OSC-MSE algorithm ($\delta = 0.8$) and corresponding RMSE. The arrow indicates the first point after the frequency transition for which the RMSE was lower than the considered threshold. The time difference between this point and the frequency transition was used to compute the estimation delay.

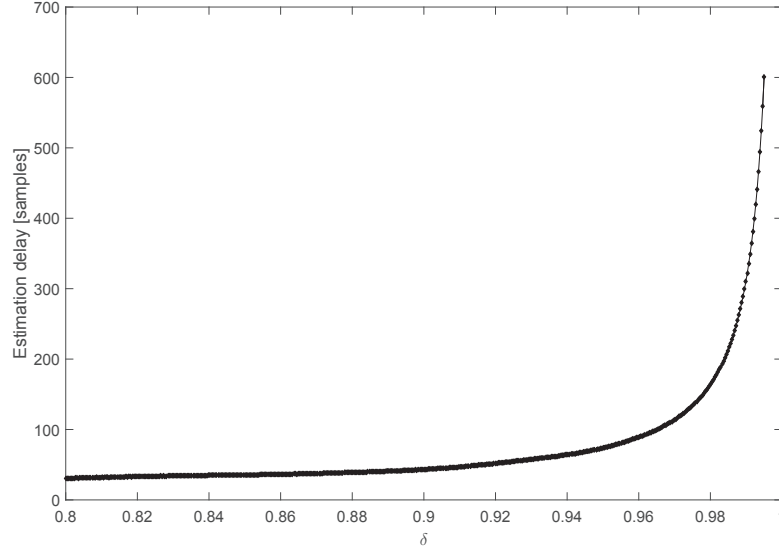


Figure 2.4: The estimation delay of the OSC-MSE algorithm as a function of δ .

Multi-signal extension

The OSC-MSE adaptive frequency tracking scheme has been extended to the multivariate case, in order to track the common frequency component present in M input signals [20, 21]. In this multi-signal extension, referred to as OSC-MSE-W, the same BP filter is used on each signal to compute individual updates and frequency estimates, similarly to the univariate case. Then, the update of the center frequency of the filter is computed as a weighted average of the individual update estimates. The computation of the weights W_m is based on the minimization of the variance of the linear combination of the individual instantaneous frequency estimates. The recursive estimates of the variances of the input signals x_m and output signals y_m ($m=1, \dots, M$) are first computed using the following equations

$$\begin{aligned} S_{ym}[n] &= \mu S_{ym}[n-1] + (1-\mu)(y_m[n] - 2\alpha[n]y_m[n-1] + y_m[n-2])^2 \\ S_{xm}[n] &= \mu S_{xm}[n-1] + (1-\mu)|x_m[n]|^2 \end{aligned} \quad (2.17)$$

with μ , a forgetting factor. The computation of the weights is given by

$$W_m[n] = \frac{S_{xm}[n]/S_{ym}[n]}{\sum_{i=1}^M S_{xi}[n]/S_{yi}[n]} \quad (2.18)$$

Finally, the global instantaneous frequency estimate is defined as

$$\omega_{global}[n] = \sum_{m=1}^M W_m[n] \alpha_m[n] \quad (2.19)$$

with $\alpha_m[n]$ ($m=1, \dots, M$), the individual frequency estimates. The scheme in Figure 2.5 shows the configuration of the multi-signal adaptive frequency tracker.

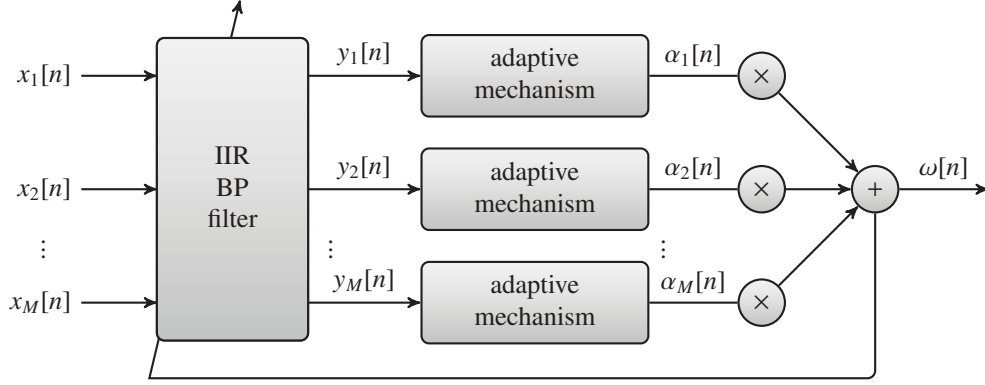


Figure 2.5: Configuration of the multi-signal adaptive frequency tracker (OSC-MSE-W).

2.3.2 SFT algorithm

This algorithm, previously described in [21, 24], was derived from the discrete oscillator based adaptive notch filter proposed by Liao [19]. In this case, the input signals is defined as:

$$x[n] = c[n] + v[n] \quad (2.20)$$

with $c[n] = A_0 e^{j\omega_0 n}$ a cisoid, with A_0 and ω_0 its complex amplitude and frequency, and $v[n]$, an additive centered complex noise. The transfer function of the time-varying single pole BP filter used in this algorithm is expressed as follows:

$$G(z; n) = \frac{1 - \beta}{1 - \beta e^{j\omega[n]} z^{-1}} \quad (2.21)$$

with $\omega[n]$, the normalized instantaneous frequency estimate and β ($0 \ll \beta < 1$), a factor related to the bandwidth of the filter. The adaptive mechanism used to update the central frequency of the BP filter described in equation 2.21 requires at each time step the minimization of a cost function, which is derived from the complex oscillator equation:

$$c[n] = e^{j\omega_0} c[n - 1] \quad (2.22)$$

By considering an input signal consisting of a complex sinusoid corrupted by a complex interference, from (2.22), the output signal $y[n]$ can be written as:

$$y[n] = \theta[n] y[n - 1] + \epsilon[n] \quad (2.23)$$

with $\epsilon[n]$, the error term and $\theta[n] = e^{j\omega[n]}$. A minimization of the MSE leads to the following expression for $\theta[n]$:

$$\theta[n] = \frac{E[y[n] \bar{y}[n - 1]]}{E[|y[n - 1]|^2]} \quad (2.24)$$

which in practice can be recursively estimated as:

$$\hat{\theta}[n] = \frac{Q[n]}{P[n]} = \frac{\delta Q[n - 1] + (1 - \delta) y[n] \bar{y}[n - 1]}{\delta P[n - 1] + (1 - \delta) |y[n - 1]|^2} \quad (2.25)$$

where the convergence rate can be adjusted with a forgetting factor δ , ($0 \ll \delta < 1$). Finally, the instantaneous frequency estimate $\omega[n]$ is defined as:

$$\omega[n] = \arg \left(\frac{\hat{\theta}[n]}{|\hat{\theta}[n]|} \right) \quad (2.26)$$

Figure 2.6 shows the amplitude and phase response of the BP filter of the SFT algorithm, for different values of β . It should be mentioned that, in practice, most of the signals are real-valued. In that case, the analytic representation of the signal, computed with the discrete Hilbert transform, is provided as input to the SFT algorithm.

Multi-signal extension

Similarly to the OSC-MSE algorithm, the SFT algorithm has been extended to the multivariate case in order to track the common frequency component present in M input signals [20, 21]. In this multi-signal extension, referred to as SFT-W algorithm, the same BP filter is applied to each signal to compute individual updates and frequency estimates, similarly to the univariate case. Then, the update of the filter central frequency is computed as a weighted average of the individual update estimates. The computation of the weights W_m is based on the minimization of the variance of the linear combination of the individual instantaneous frequency estimates. The recursive estimates of the variances of inputs signals x_m and output signals y_m ($m=1,\dots,M$) are first computed using the following equations

$$\begin{aligned} S_{ym}[n] &= \delta S_{ym}[n-1] + (1-\delta)|y_m[n]|^2 \\ S_{xm}[n] &= \delta S_{xm}[n-1] + (1-\delta)|x_m[n]|^2 \end{aligned} \quad (2.27)$$

The weights are computed as:

$$W_m[n] = \frac{S_{xm}[n]/S_{ym}[n]}{\sum_{i=1}^M S_{xi}[n]/S_{yi}[n]} \quad (2.28)$$

Finally, the global instantaneous frequency estimate is defined as

$$\omega[n] = \sum_{m=1}^M W_m[n]\omega_m[n] \quad (2.29)$$

2.3.3 Inclusion of signal quality in adaptive frequency tracking schemes

Biomedical signals recorded in a non-invasive way can contain artifacts, rendering their analysis very challenging. When estimating the instantaneous frequency with the OSC-MSE/SFT algorithms, sudden drops in signal quality usually result in large displacements of the BP central frequency. As a consequence, the algorithm takes some time after the end of the perturbation to re-converge to the frequency of the main oscillatory component. In the framework of this thesis, which focuses on the processing of photoplethysmographic signals, some modifications of the OSC-MSE-W and SFT-W algorithms were proposed in order to incorporate the signal quality into these multi-signal adaptive frequency tracking algorithms and therefore improve the robustness of frequency estimation. The next two paragraphs describe two modifications of the adaptive mechanism developed for the two following cases: 1) the signal quality is available and 2) the signal quality is not available.

Available signal quality

In some situations, signal quality information is available or can be assessed with a specific method, for instance by using the information provided by an additional sensor (e.g. an accelerometer). Quality measures for biomedical signals are usually named signal quality indexes (SQIs). In the framework of this thesis, a modification of the OSC-MSE-W algorithm was proposed in order to incorporate a SQI into this multi-signal adaptive frequency tracking scheme.

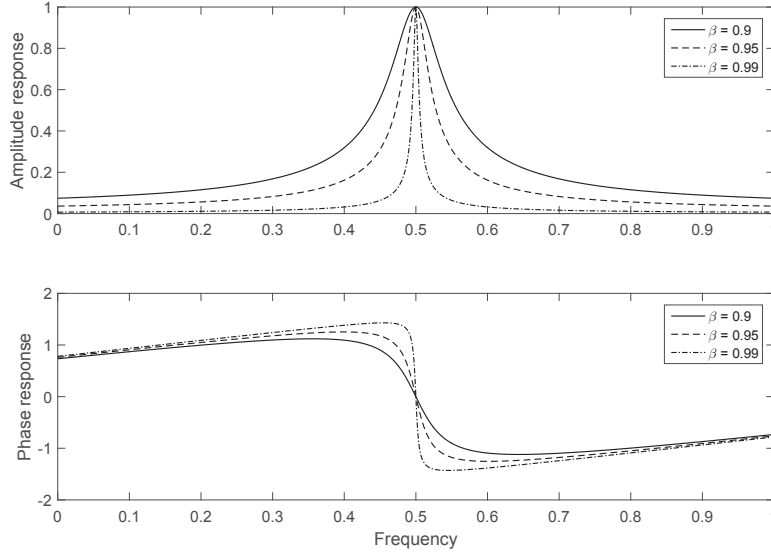


Figure 2.6: Amplitude and phase response of the BP filter used in the SFT algorithm, for different values of β .

This modification aimed at increasing the robustness of frequency estimation in presence of artifacts and at providing an indication about the reliability of the current frequency estimate. This algorithm is referred to as OSC-MSE-W-SQI. It is assumed here that the SQI of a signal is a vector of the same length as the signal, taking the value of one when the signal quality is acceptable, zero otherwise. We decided to include a SQI in the adaptive frequency tracking scheme in order to freeze filter adaptation during bad quality epochs, via the forgetting factors δ and μ . These scalars were replaced by vectors, in order to have both time-dependent and signal-dependent forgetting factors $\delta_m[n]$ and $\mu_m[n]$. In this case, equations 2.15 become:

$$\begin{aligned} Q_m[n] &= \delta_m[n]Q_m[n-1] + (1 - \delta_m[n])y_m[n-1](y[n] + y_m[n-2]) \\ P_m[n] &= \delta_m[n]P_m[n-1] + (1 - \delta_m[n])(y_m[n-1])^2 \end{aligned} \quad (2.30)$$

and equations 2.17 become:

$$\begin{aligned} S_{ym}[n] &= \mu_m[n]S_{ym}[n-1] + (1 - \mu_m[n])|y_m[n] - 2\alpha[n]y_m[n-1] + y_m[n-2]|^2 \\ S_{xm}[n] &= \mu_m[n]S_{xm}[n-1] + (1 - \mu_m[n])|x_m[n]|^2 \end{aligned} \quad (2.31)$$

and the forgetting factors were defined according to the SQI values:

$$\delta_m[n] = \begin{cases} 1, & \text{if } \text{SQI}_m[n] = 0 \\ \delta \text{ as defined previously with } (0 \ll \delta < 1), & \text{otherwise} \end{cases} \quad (2.32)$$

$$\mu_m[n] = \begin{cases} 1, & \text{if } \text{SQI}_m[n] = 0 \\ \mu \text{ as defined previously with } (0 \ll \mu < 1), & \text{otherwise} \end{cases} \quad (2.33)$$

This OSC-MSE-W-SQI algorithm was tested on synthetic signals, consisting of two 3000-sample long sinusoids at $\omega_0 = 0.2\pi$ with uniformly distributed random phases and with additive Gaussian white noise at various SNR levels ranging between 0 and 30 dB (with one dB step).

In order to simulate bad-quality segments, a few local perturbations were added to these signals. More precisely, for seven manually selected 40-sample long epochs, the SNR was set to -12 dB and the SQI was set to zero for the corresponding samples. The instantaneous frequency was then computed using the OSC-MSE-W and OSC-MSE-W-SQI algorithms and bias and variance analysis was performed in order to evaluate the performance of the two algorithms on the aforementioned synthetic signals. The parameters β , δ and μ were set to 0.9. For each SNR value, 1000 runs were performed and frequency bias, variance and average absolute error (AAE) were computed over the last 2000 samples. Results are shown in Figure 2.7. It can be noted that both bias and variance were lower for the SQI-modified algorithm than for the original algorithm. Indeed, in this example, the variance is reduced, on average, by 97% and the AAE by 73%. As expected, the OSC-MSE-W-SQI results in an unbiased estimate at high SNR values, despite the presence of the bad-quality segments. Figure 2.8 shows an example of these signals, as well as their frequency estimates using the original and the modified algorithms. It can be noticed that the OSC-MSE-W-SQI algorithm is more robust to the local disturbances, especially when the two signals are simultaneously polluted by noise, which is often the case with biomedical signals (e.g. when patient movements affect all signals). In such cases, the oscillation of interest is lost from the signal for a few samples, perturbing the original algorithm, which will take time to re-converge to the true frequency. It should be emphasized that this OSC-MSE-W-SQI algorithm requires an accurate input SQI in order to be efficient, which is itself a challenging task.

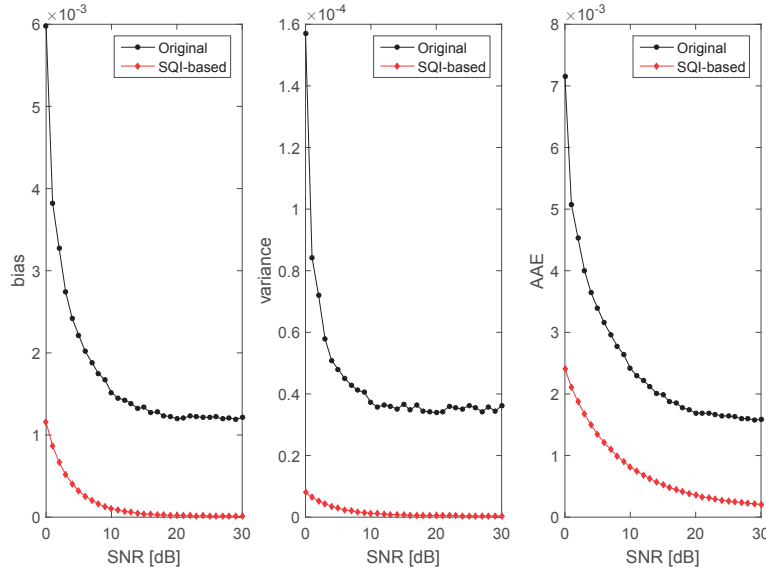


Figure 2.7: Bias, variance and AAE for the OSC-MSE-W and OSC-MSE-W-SQI algorithms. The inputs signals consisted of two 3000-sample long sinusoids at $\omega_0 = 0.2\pi$ with uniformly distributed random phases and with additive Gaussian white noise at various SNR levels ranging between 0 and 30 dB (with one dB step). In order to simulate bad-quality segments, seven manually selected local 40-sample long perturbations (SNR of -12 dB) were added to each signal.

Signal quality is not available

Sometimes the SQI of the studied signal is not directly available. In such cases, a method to estimate the quality directly from the waveform is needed. The following observation can

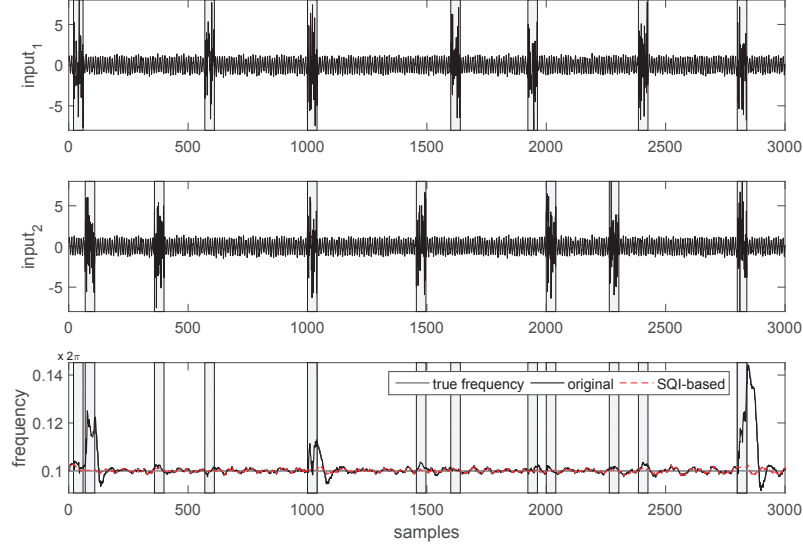


Figure 2.8: Example of two synthetic input signals. The SNR is 15 dB in this example and -12 dB during the simulated bad-quality regions (shaded zones). The frequency estimates using both OSC-MSE-W and OSC-MSE-W-SQI algorithms are shown (bottom).

be made from the spectrum of a signal of interest: the more the power is concentrated around a single frequency peak, the easier frequency estimation is. We decided to use a parameter called the spectral purity index (SPI) to quantify how well the signal is locally defined by a single frequency. The SPI is described in detail in the next section (Section 2.4). This new SPI-based adaptive mechanism was applied to the SFT-W algorithm. In the original version of the algorithm, δ is a fixed forgetting factor that should be chosen as a trade-off between convergence speed and estimation variance. The parameter β is related to filter bandwidth, with β closer to one indicating a narrower filter. Adaptive mechanisms for δ and β were developed based on the following considerations. Ideally, δ should be smaller when the signals are of good quality, in order to guarantee a faster tracking. On the other hand, a larger δ value is required to ensure stability when there is a reduction in signal quality. Regarding the adaptation of β , the estimation variance can be reduced by using a β closer to one. However, when the quality of the signals is not high enough, smaller values of β ensure the tracking of an oscillation. Then, δ and β were defined as linear saturated functions of the average SPI (averaged across the different channels if more than one input signal is provided). These functions are shown in Figure 2.9.

The modified algorithm was evaluated on synthetic signals. In order to make comparisons possible, the instantaneous frequency estimation was performed with both the original SFT-W and the modified SFT-W-SPI algorithm. For the original version β and δ were set to 0.99 and 0.95, respectively, as these values correspond to the ones used in the modified algorithm when the signal quality is very good ($\text{SPI} > 0.9$). Two different types of input signal were used. In the first case, the input signal consisted of a 3000-sample long sinusoid at $\omega_0 = 0.2\pi$ with an uniformly distributed random phase. Gaussian white noise was added at various SNR levels ranging from 0 to 30 dB (with one dB step). For each SNR value, 1000 runs were performed and frequency bias, variance and AAE were computed over the last 2000 samples. In the second case, segments with a poor signal quality were simulated by lowering the SNR to -5 dB for two epochs of 200 samples.

The results for the two cases are shown in Figures 2.10 and 2.11, respectively. It can be noted in Figure 2.10 that the bias was larger at low SNR values for the SFT-W-SPI, whereas variance and AAE were smaller. For SNR values above approximately 12 dB, the performance of the two algorithms are very similar. This was expected as δ and β had the same value in both algorithms when the SPI was higher than 0.9. It can be noted in Figure 2.11 that, in the presence of bad-quality segments, the variance and the AAE are smaller for SFT-W-SPI. On average, the AAE was reduced by 51% for this example. Figure 2.12 shows an example of the input signal and the resulting SPI, δ , β and corresponding frequency estimates. It can be noted that the SPI decreases in the two regions with low SNR, resulting in an δ increase and a β decrease. As a result, it can be noted that the modified version of the algorithm is less affected by the simulated bad-quality segments. This property is particularly interesting for real-time applications. Indeed, it allows to keep a good convergence rate while increasing the stability during segments containing artifacts without introducing a trade-off between stability and convergence rate when dealing with signals of variable quality.

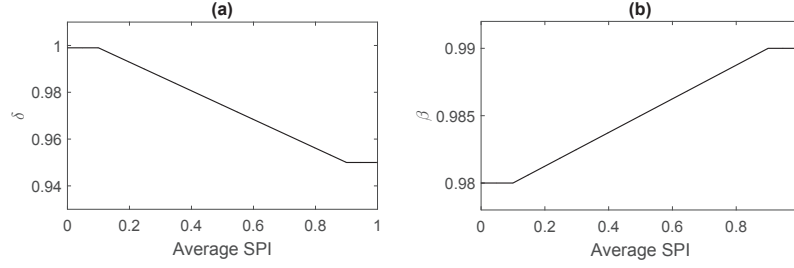


Figure 2.9: Adaptive parameters. (a) δ as a function of the average SPI; (b) β as a function of the average SPI.

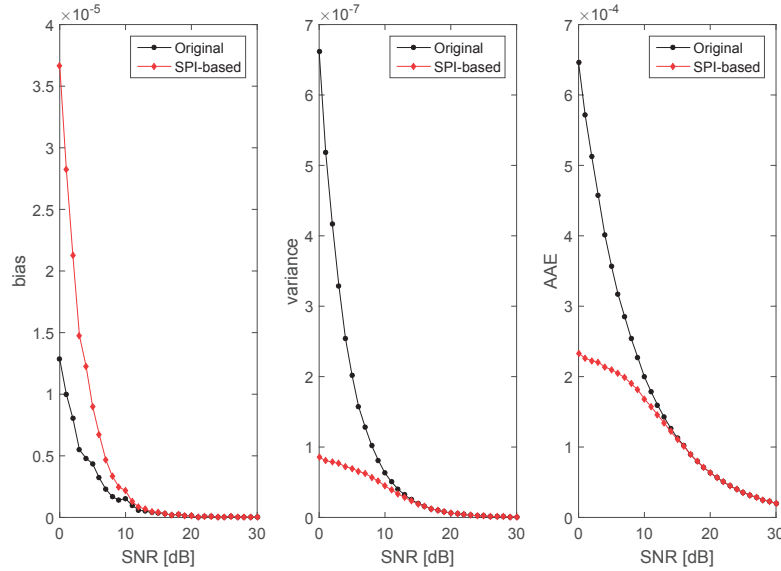


Figure 2.10: Bias, variance and AAE of frequency estimation with STF-W and SFT-W-SPI algorithms averaged over 1000 runs for each tested SNR value, and for a synthetic signal consisting of a sinusoid at $\omega=0.2\pi$.

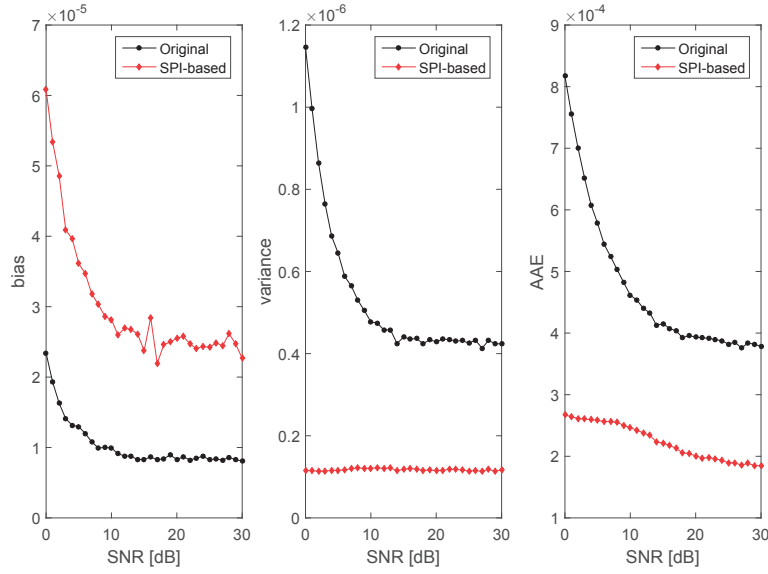


Figure 2.11: Bias, variance and AAE of frequency estimation with STF-W and SFT-W-SPI algorithms averaged over 1000 runs for each SNR value, and for the synthetic signal consisting of a sinusoid at $\omega=0.2\pi$ and containing simulated bad-quality segments.

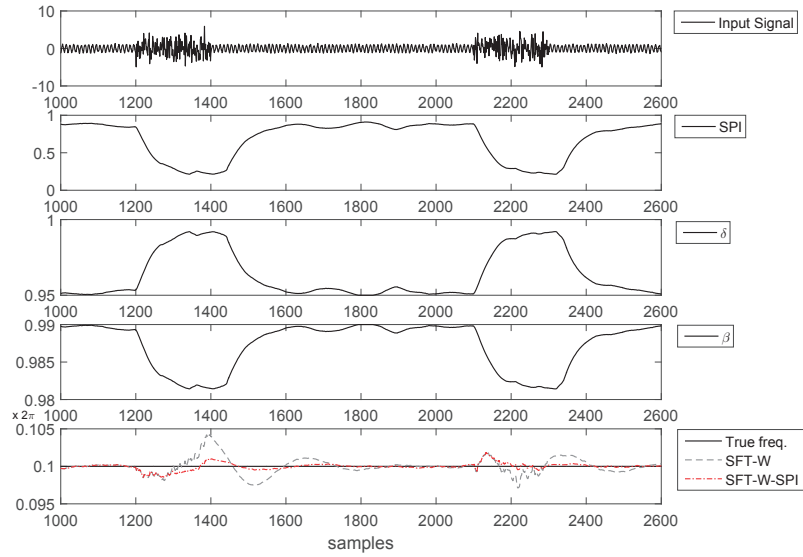


Figure 2.12: Frequency estimation using the SFT-W-SPI algorithm on a synthetic signal. The input signal consist of a noisy sinusoid (SNR = 15 dB) with simulated bad-quality epochs (SNR = -5 dB). The time evolutions of the SPI, δ and β are shown, as well as the resulting frequency estimates.

2.4 Spectral purity index

The adaptive frequency tracking algorithms require a narrow enough signal bandwidth so that the oscillatory component of interest can be properly isolated. In addition, the bandwidth of any biomedical signal is itself a potentially useful information, as it can be related to signal quality or morphological changes possibly related to physical state changes. The spectral purity index is a heuristic parameter designed to provide information about signal bandwidth. This measure, which was originally developed in the context of the analysis of electroencephalogram signals [9, 25], ranges between zero and one and indicates how well the signal of interest can be described by a single frequency. It is defined as the running squared second-order spectral moment divided by the product of the running total power and fourth-order spectral moment:

$$\text{SPI}(n) = \frac{\bar{\omega}_2^2(n)}{\bar{\omega}_0(n)\bar{\omega}_4(n)} \quad (2.34)$$

with $\bar{\omega}_n$ the n th-order spectral moment, defined by:

$$\bar{\omega}_n = \int_{-\pi}^{\pi} \omega^n S_x(e^{j\omega}) d\omega \quad (2.35)$$

where $S_x(e^{j\omega})$ represents the power spectrum of the signal. In the present thesis, the spectral moments $\bar{\omega}_n$ were estimated in the time domain, as proposed by [9]. The following difference equations can be used to estimate the first and second derivatives:

$$x^{(1)}[n] = x[n] - x[n-1] \quad (2.36)$$

$$x^{(2)}[n] = x[n+1] - 2x[n] + x[n-1] \quad (2.37)$$

Then, from these derivatives, spectral moments can be estimated as:

$$\hat{\omega}_i \approx \frac{2\pi}{N} \sum_{n=0}^{N-1} (x^{(i/2)}[n])^2 \quad (2.38)$$

with $i = 0, 2, 4$ and N the length of the signal. In addition, the SPI can be measured recursively using a sliding window of length L . The expression for the $\text{SPI}[n]$ is then written as:

$$\text{SPI}[n] = \frac{(\sum_{k=n-L+1}^n (x^{(1)}[k])^2)^2}{\sum_{k=n-L+1}^n (x^{(0)}[k])^2 \sum_{k=n-L+1}^n (x^{(2)}[k])^2} \quad (2.39)$$

Depending on the application, one may want to smooth this instantaneous SPI estimate in order to focus more on long-term variations. For this purpose, an optional forgetting factor λ was incorporated into the SPI estimation scheme as follows:

$$\text{SPI}_{\text{smooth}}[n] = \lambda \cdot \text{SPI}_{\text{smooth}}[n-1] + (1 - \lambda) \cdot \text{SPI}[n] \quad (2.40)$$

Figure 2.13 shows the SPI of a synthetic signal in three different regimes: a sinusoid at frequency $\omega_0 = 0.2\pi$, a sum of two sinusoid at frequencies $\omega_0 = 0.2\pi$ and $\omega_1 = 0.08\pi$, and the sum of the same two sinusoids with an additive Gaussian noise (SNR = 0 dB). As expected, the SPI is equal to one during the first regime, consisting of a pure sinusoid. The SPI decreases during the second regime because the signal is composed of a sum of two sinusoids at different frequencies. During the third regime, the SPI decreases more, due to the presence of the white Gaussian noise. In the framework of this thesis, this SPI measure was used in three applications, namely the evaluation of PPG signal quality (see Chapter 4), the assessment of PPG signal organization during arrhythmias (see Chapter 5) and the detection of ventricular tachycardia on ECG signals (see Chapter 6).

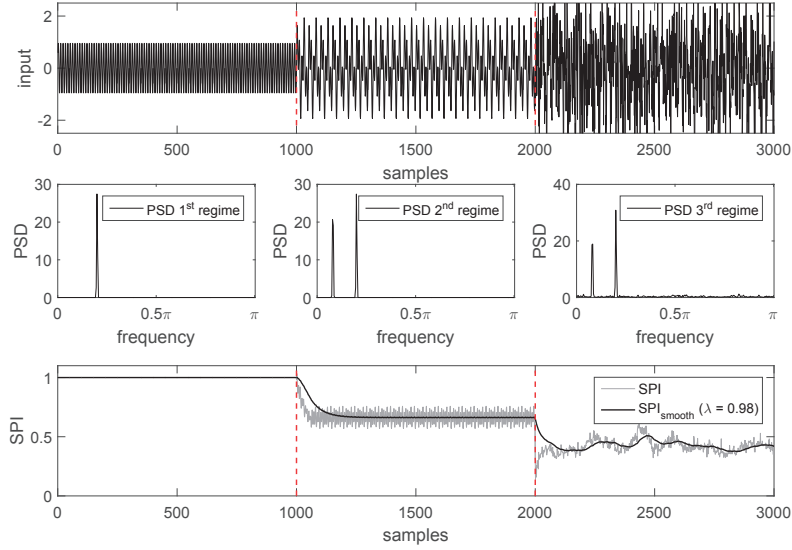


Figure 2.13: The SPI of a synthetic signal with three different regimes, using $L = 60$ and $\lambda = 0.98$. Original signal (top), PSD for each regime (middle) and instantaneous SPI (below).

2.5 Conclusion

The concept of instantaneous frequency was discussed in this chapter. Two adaptive frequency tracking algorithms were presented, one designed to work with real-valued signals (OSC-MSE) and one designed to work with complex-valued signals (SFT). Their adaptive mechanisms, based on the (complex) oscillator equation, are designed to maximize the oscillatory behavior, resulting in scale-independent algorithms able to extract an oscillation with low amplitude compared to the noise level. The multi-signal extensions of these algorithms were also presented. In these multi-input configurations, the common frequency component present in the several input signals can be extracted by filtering all the inputs with the same BP filter and weighting the contribution of each input signal. It should be mentioned that these two algorithms can be used for the similar purposes. However, we noted that, in some cases, one or the other had some advantages and led to better results. In addition to the fact that the SFT is applied to the analytic representation when the input is real-valued, the difference in the filter responses also explain the difference in the obtained results. More specifically, we observed that the real version has worse performance for low normalized frequencies due to the presence of a zero in the transfer function at normalized frequency zero.

Two extensions of these adaptive frequency tracking algorithms were developed in the framework of this thesis in order to take the signal quality into account. In one case, we assumed that a binary SQI was available and used it to freeze filter adaptation during bad-quality segments. In the other case, we proposed to use the SPI to have a real-time continuous estimation of the quality of the oscillatory waveform and δ and β were defined as saturated linear function of the SPI.

We believe that these adaptive frequency tracking tools have some advantages over the Fourier-based methods. Indeed, Fourier-based methods are not optimal for real-time processing as they induce a delay of half the length of the temporal window used to calculate the spectrum. Moreover, Fourier-based methods require a peak-selection process, while the presented algorithms

are more straightforward. Finally, these adaptive frequency tracking algorithms offer an elegant solution to combine multiple signals, that do not even have to be in phase, and therefore increase the robustness of frequency estimation in most cases. For all these reasons, they were largely used in different applications presented in this thesis, namely 1) to perform robust heart rate estimation from PPG signal while subjects are exercising, 2) to perform heart rate estimation from PPG and ABP in the ICU in order to suppress false arrhythmia alarms, 3) to quantify the level of organization in PPG signal during different types of heart rhythms and 4) to perform contactless video-based heart rate estimation using iPPG.

Part II

Processing of Photoplethysmographic Signals

Introduction to Photoplethysmography

3

Photoplethysmography (PPG) is a noninvasive optical technique used to measure blood volume changes in the microvascular bed of tissue [6]. PPG technology has a wide range of applications, including, inter alia, the monitoring of various cardiovascular parameters (heart rate, blood oxygen saturation, blood pressure, cardiac output and respiration), the assessment of vascular function and the evaluation of autonomic activity through heart rate variability. This technique has the advantage of being low-cost and integrable into wearable devices. For all these reasons, there has been a surge of interest in PPG in recent years.

The sensor system is composed of a light source, usually a light-emitting diode (LED), and a photodetector. As shown in Figure 3.1, two operational configurations exist: the transmission mode, for which the biological tissue is placed between the illumination source and the photodetector and the reflection mode, for which the illumination source and the photodetector are placed side-by-side.

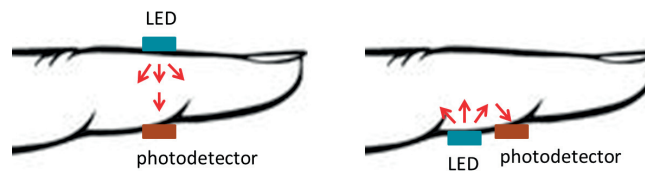


Figure 3.1: The two possible operational configurations of a PPG sensor system: the transmission mode (left) and the reflectance mode (right).

Regarding the wavelength used for the illumination source, different possibilities have been investigated. In [26], it was shown that the origin of the pulsatile AC component of the PPG signal depends on the light wavelength. Using an optic probe positioned on the left arm, the authors concluded that for green light (at 560 nm), the pulsations in the ascending arterioles were the source of the AC component while at an infrared (IR) wavelength (940 nm) the observed AC component was related to the pulsations occurring deeper in the tissue volume in larger vessels. The underlying explanation is the presence of a peak in the haemoglobin absorption spectrum at the 560 nm wavelength. Therefore, the intensity in the deeper layers of the skin is considerably reduced at this wavelength [27]. This is probably the reason why, recently, the PPG devices operating at green wavelength are becoming more popular than the ones operating at red/IR wavelengths [28–30].

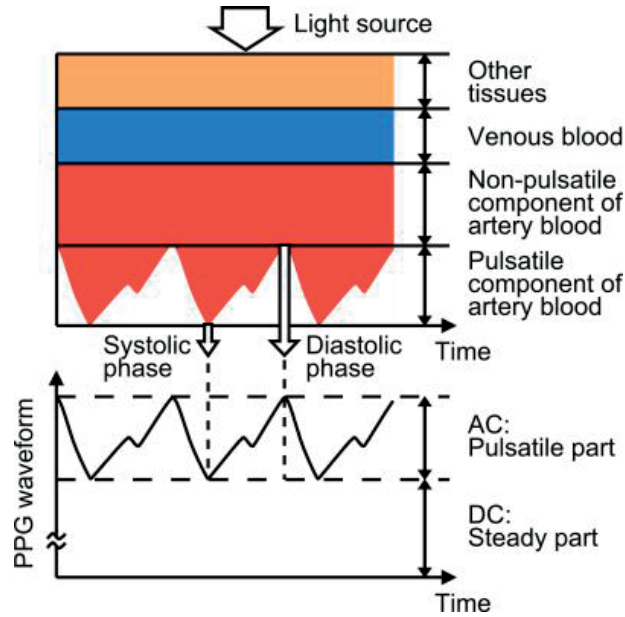


Figure 3.2: Variation in light attenuation by tissue. Image from [28], © 2016 IEEE.

In addition to the aforementioned pulsatile “AC” component, which reflects the cardiac synchronous changes in the blood volume with each heartbeat, the PPG waveform also comprises a slowly varying “DC” baseline. Figure 3.2 illustrates the origins of the AC and DC components of the PPG waveform. It is important to mention that the PPG waveform is usually inverted. The DC component reflects the influence of respiration, sympathetic nervous system activity and thermoregulation [6]. Two phases can be distinguished in an AC-PPG pulse: the anacrotic phase, corresponding to the rising edge of the pulse and mainly related to systole, and the catacrotic phase, corresponding to the falling edge of the pulse and related to the diastole and wave reflections from the periphery. A dicrotic notch is usually observed in the catacrotic phase. Figure 3.3 shows a PPG waveform below the corresponding ECG segment. Different pulse landmarks are highlighted in this illustration, namely the pulse foot, the pulse peak and the amplitude. The pulse transit time to the foot of the pulse (PTT_f) and the pulse transit time to the peak of the pulse (PTT_p), computed using the ECG R-waves, are also shown. Different sites of the body periphery can be used to record the PPG signal, the more popular being fingers, toes and ears. Figure 3.4 shows examples of pulse shapes for PPG waveforms acquired at different peripheral locations on two healthy subjects. The pulse wave characteristics depend on various factors. In [31], the evolution of the pulse shape characteristics with aging was studied, for PPG recorded at the fingers, the toes and the ears. The authors observed a significant overall elongation of the systolic rising edge, which was attributed to changes such as resistance and compliance occurring in arterial properties. A damping of the dicrotic notch was also reported, which can be partly explained by age-related changes in pulse velocity. In another study by the same authors, the similarity between left-body-side and right-body-side pulse wave characteristics was studied [32]. For this purpose, PPG signals were acquired simultaneously at the ears, the fingers and the toes and the similarity between pulse shapes was assessed using root mean square error and cross-correlation. Pulses from the left and right sides were highly correlated for normal subjects. However, some differences were observed in vascular patients suffering from asymmetrical diseases.

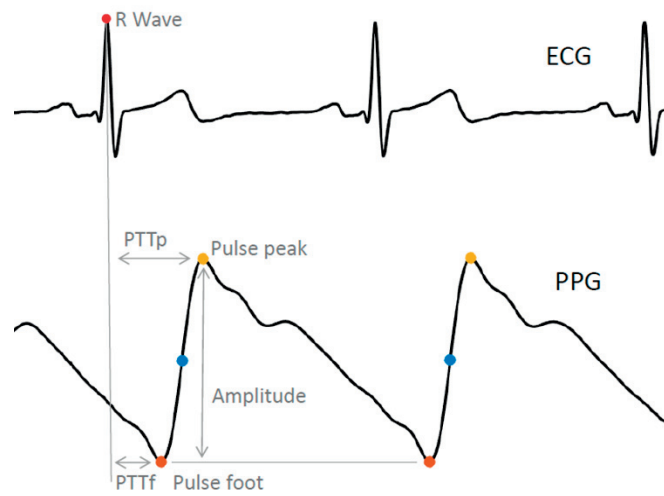


Figure 3.3: PPG waveform and pulse landmarks frequently used. PTTf is the pulse transit time to the foot of the pulse and PTTp is the pulse transit time to the peak of the pulse.

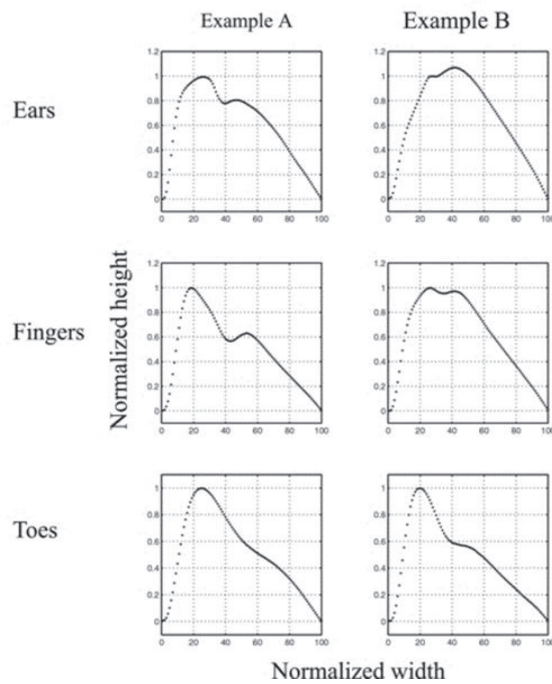


Figure 3.4: Example of waveforms acquired in two healthy subjects at different body locations. Figure from [6].

In the last years, most efforts have been concentrated to provide low-cost wearable PPG systems, comfortable to wear and suitable for long-term recordings. However, various challenges linked to data quality have arisen. In [4], the following three main limitations of PPG were pointed out: spot measurement, contact measurement and motion artifact corruption. Spot measurement is the fact that a probe monitors a single site, resulting in a very localized picture of the dynamic changes of blood volume. Contact measurement means that the PPG sensor has to be tightly attached to the skin. Indeed, it has been shown that the contact force has to be precisely controlled in order to obtain high quality data [33–36]. Finally, motion artifacts are an important problem when dealing with PPG signals. In order to lower their impact, an accelerometer sensor is often embedded in the device, providing a reference for motion. This aspect will be discussed more into detail in Chapter 4.

Two main applications of PPG are targeted in the framework of this thesis, namely the estimation of heart rate and the detection of arrhythmias. A considerable number of approaches have already been proposed to estimate heart rate from the PPG signals. However, the number of PPG devices commercialized for out-of-hospital use has exploded these last years and being able to provide reliable heart rate estimates during all kind of daily life activities is a current challenge. On the other hand, the potential of PPG for the detection of arrhythmias has not been extensively studied yet. We believe that this field will be explored in depth in the next years. Indeed, PPG wearable devices have a huge potential as a screening tools as they allow to perform long-term monitoring out of the hospital. The next three chapters are dedicated to 1) the development of a processing scheme to perform robust heart rate estimation from PPG signals when the subjects are performing physical exercise (Chapter 4), 2) the detection of atrial fibrillation using a wrist-type PPG device (Chapter 5) and 3) the development of a multimodal approach to reduce the number of false alarms in the ICU, which includes the information from the PPG signals (Chapter 6).

Robust Heart Rate Estimation Based on Adaptive Filtering

4

The purpose of this chapter is to introduce a new methodology for robust heart rate estimation during physical exercise using wrist-type PPG signals corrupted by strong motion artifacts.

4.1 Introduction

In recent years, wearable pulse rate sensors based on PPG technology have emerged as very promising tools to monitor HR outside of the hospital. Indeed, these biosensors can be worn for all kind of daily activities, including sport sessions and are useful to evaluate the physical condition of subjects and prevent injuries [2]. The commonly used methods to estimate HR from the AC component of PPG waveforms are Fourier-based methods. However, these techniques may not be an optimal choice because of the non-stationary nature of PPG signals. In addition, the presence of motion artifacts (MA) can strongly affect the quality of the waveforms and make them unusable, as illustrated in Figure 4.1. Therefore, hardware and software improvements are still required to fully exploit the potential of wearable PPG sensors. Different approaches have been investigated so far to identify and remove MA from PPG signals. Some of these methods involve various types of frequency-domain data processing, including smoothed pseudo Wigner-Ville distribution [37], frequency-domain independent component analysis [38] and Fourier series analysis on a cycle-by-cycle basis [39]. Another possible course of action to reduce MA in PPG waveforms is adaptive filtering. Different variants of least-mean-squares (LMS) filters were tested using either accelerometer (ACC) signals as noise reference [40] or a synthetic noise [41]. The aforementioned strategies for the attenuation of MA have been compared by Narahariseti *et al.* [42]. The authors concluded that, in the absence of extra hardware such as an accelerometer, cycle-by-cycle Fourier analysis and singular value decomposition were the most efficient methods to remove MA.

The removal of MA is particularly important for applications involving HR monitoring during sport sessions. Recently, some studies have reported complete schemes for HR estimation using PPG and ACC signals during physical exercise. An approach based on signal decomposition for denoising, sparse signal reconstruction, and spectral peak-tracking (TROIKA) was developed by Zhang *et al.* [43]. In this HR estimation scheme, a high-resolution spectrum is obtained by means of a sparse signal reconstruction and the spectral peak-tracking step comprises an initialization, followed by peak selection and verification. Another method based on joint sparse spectrum reconstruction using a multiple measurement vector model, followed by spectral subtraction and spectral peak-tracking (JOSS) was developed by Zhang [44]. In this study, a sparsity constraint is used on spectral coefficients to identify and remove the spectral peaks corresponding to MA

on the PPG spectra. In the two aforementioned studies, the success of the subsequent spectral peak tracking step strongly relies on a proper initialization. In order to reduce the importance of this initialization step, Lakshminarasimha Murthy *et al.* developed a framework using a multiple initialization scheme for spectral peak tracking (MISPT), in which MA removal is first performed using adaptive noise cancellation and then, the multiple HR trajectories are combined to provide an accurate HR estimation [45]. Another approach based on empirical mode decomposition was proposed by Khan *et al.*, with a recursive least-squares filter used to remove MA. Finally, spectral-peak tracking is performed on the different reconstructed and original PPG signals [46]. In another study, Sun and Zhang introduced a framework (SPECTRAP) involving a new spectrum subtraction algorithm based on asymmetric least-squares to remove MA from the PPG spectrum [47]. Then, a spectral peak-tracking method based on Bayesian decision theory is used to extract HR estimates. The performance of the above mentioned HR estimation schemes [43–46] was evaluated on the same database, made available by the authors of [43].

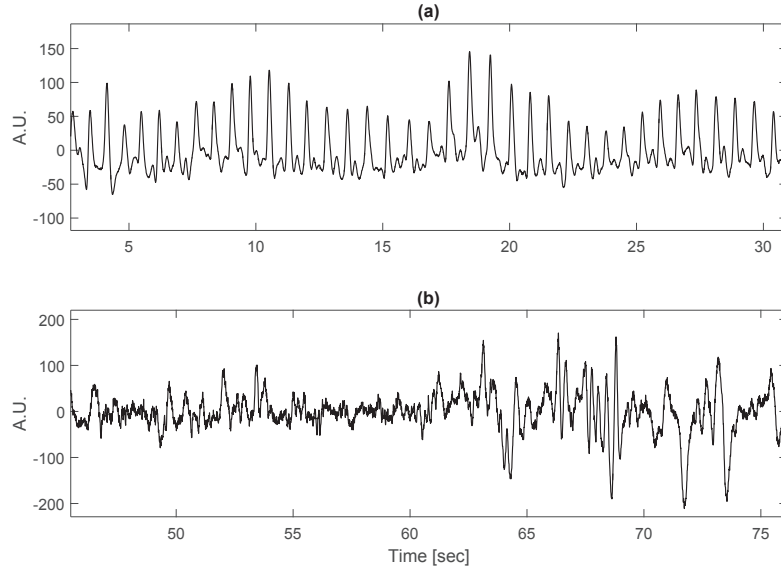


Figure 4.1: Example to illustrate how motion artifacts can corrupt the PPG waveforms. (a) Segment of good-quality PPG signal; (b) Segment of PPG signal corrupted by motion artifacts.

4.1.1 Motivation of the proposed study

The AC component of PPG has been proven to be reliable for HR estimation. However, as mentioned previously, PPG waveforms are easily affected by motion artifacts. Therefore, the development of new algorithms to increase the reliability of the estimated HR when the subjects are performing daily-life activities is a current challenge. Two main points should be taken into consideration in the development of such algorithms. Firstly, the computational cost should not be high if the aim is to embed them in a wearable device. Secondly, as the goal is to provide instantaneous HR values, the estimation delay should be as short as possible.

The purpose of the study presented in this chapter is to develop a new approach for HR estimation using wrist-type PPG signals corrupted by strong MA. Although existing methods lead to good results, it should be noted that all of them are based on a block-wise implementation, which inevitably introduces half block-length delay in the processing flow. Moreover, the spectral-peak tracking step often relies on many assumptions and a considerable number of parameters have to

be precisely tuned in order to achieve satisfying results. The method we introduce in this chapter is straightforward and mainly operates in the time domain. MA reduction and frequency tracking are achieved by two consecutive adaptive filters, on an almost real-time basis. Moreover, the method does not require any a priori knowledge about the signals. This work was the object of two publications [48, 49].

4.2 Methods

4.2.1 Data

The IEEE Signal Processing Cup 2015 database is composed of 22 recordings and was made available by the authors of [43]. The details about age, sex and physical condition of the subjects are provided in Table 4.1. For each subject, the following waveforms were recorded simultaneously: three-axis acceleration (ACC), one channel ECG and two-channel PPG. Each waveform was sampled at 125 Hz. The two pulse oxymeters (using green LEDs) as well as the accelerometer sensor were embedded into a wristband. The ECG signal, recorded from the subject's chest, was available only for the first 12 records. However, for every subject, a ground-truth HR (BPM_{ref}), derived from ECG, was provided. BPM_{ref} was defined as the average HR value in every 8-second time window (6-second overlap). Each recording lasted approximately 5 minutes. More precisely, it corresponded to an average of 140.73 ± 15.30 time windows (or BPM_{ref} values), per subject. The different kinds of physical activities that were performed by the subjects are specified in Table 4.1. For the first 12 records (dataset-1), the subjects were required to run (R) on a treadmill at different speeds according to the following protocol: 0.5 min at 1-2 km/h, 1 min at 6-8 km/h, 1 min at 12-15 km/h, 1 min at 6-8 km/h, 1 min at 12-15 km/h and 0.5 min at 1-2 km/h. The second part of the database (dataset-2) is composed of records during which the subjects were asked to perform different forearm and upper arm movements (T). T01 corresponds to various exercises such as hand shakes, stretches, pushes (common in arm rehabilitation), running, jumps and push-ups, whereas T02 corresponds to intensive forearm and upper arm movements (e.g. boxing).

Table 4.1: Database description. BP: blood pressure.

	Record number	Sex	Age	Healthy?	Activity Type
Dataset-1	1-12	M	[18-35]	Y	R
	13,14,19	M	[18-35]	Y	T01
Dataset-2	15,16,17,18,20,21	M	[18-35]	Y	T02
	22	F	58	Abnormal HR and BP	T01

4.2.2 Adaptive motion artifact reduction

Adaptive filters, i.e. digital filters with adaptive coefficients, have a wide range of applications. The general configuration of adaptive filters is shown in Figure 4.2. In this figure, n is the time index, $x[n]$ is the input signal, $y[n]$ denotes the adaptive filter output and $d[n]$ is the desired signal. The error signal $e[n]$ is computed as $(d[n] - y[n])$. The update of the filter coefficients is achieved through the adaptation algorithm, which uses the error signal in an objective function minimized to match the desired signal $d[n]$ [50]. There are two main optimization methods; the Newton ones and gradient descent ones. This paragraph focuses on one application of adaptive filtering, namely adaptive noise cancellation.

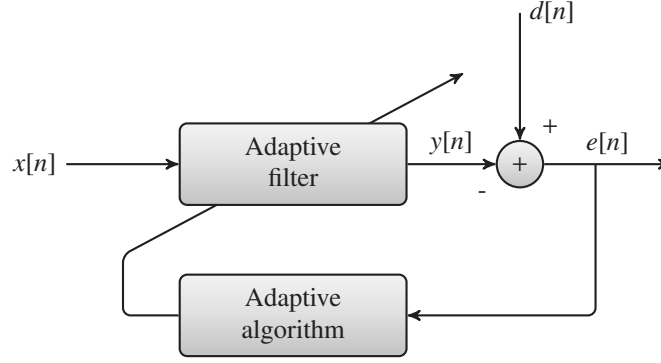


Figure 4.2: General configuration of adaptive filters.

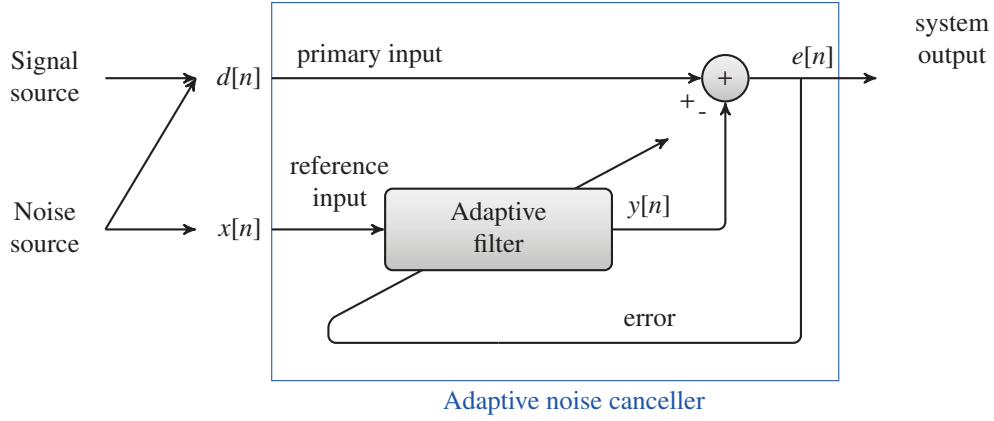


Figure 4.3: Concept of adaptive noise cancelling.

Adaptive noise cancelling (ANC) is a method to estimate signals corrupted by additive noise or interference. Two inputs are needed; the “primary” input and the “reference” input. The primary input contains the signal corrupted by noise and the reference input is derived from sensors located at points in the noise field where the signal is weak or undetectable. As illustrated in Figure 4.3, these two signals are provided to the adaptive noise canceler, where the reference input is adaptively filtered and then subtracted from the primary input to retrieve the signal estimate [12, 51]. ANC algorithms have been previously used in the context of PPG signals, using either accelerometer signals as noise reference [40], or a synthetic noise signal generated from the MA-corrupted PPG signal [41], or a signal from an optoelectronic sensor [52]. In the present study, based on the hypothesis that MA interferences in the PPG and ACC waveforms are correlated, we decided to use a normalized least-mean-square (NLMS) algorithm with the ACC signals as noise reference. This algorithm is widely used in adaptive filtering due to its computational simplicity [50]. In this well known algorithm, an input signal $x[n]$ is provided as well as a desired signal $d[n]$. When the aim is to perform ANC, $x[n]$ is the reference input and $d[n]$ is the primary input (see Figure 4.3). The filter output $y[n]$ minimizes the least-mean-squares error, and the error signal $e[n]$ is used to update the filter coefficients. The output error is given by:

$$e[n] = d[n] - \mathbf{x}^T[n]\mathbf{w}[n] \quad (4.1)$$

with $\mathbf{w}[n]$ the filter coefficient vector, and $\mathbf{x}[n]$ the input vector composed of samples of $x[n]$.

The equation for the weight-vector update is given by:

$$\mathbf{w}[n+1] = \mathbf{w}[n] + \frac{\mu e[n] \mathbf{x}[n]}{\gamma + \mathbf{x}^T[n] \mathbf{x}[n]} \quad (4.2)$$

with γ , a small constant to avoid large step size when $\mathbf{x}^T[n] \mathbf{x}[n]$ becomes small and μ a fixed convergence factor, $0 < \mu \ll 2$.

The ACC and PPG waveforms were first re-sampled at 35 Hz prior to adaptive MA reduction using the NLMS algorithm. Importantly, as the contribution of each ACC-axis to the deterioration of the PPG waveforms was not known, adaptive MA reduction was performed for the six possible ACC-PPG channel pairs. For this purpose, the corrupted PPG signals were defined as primary input signals $d[n]$ and ACC signals were defined as reference input signals $x[n]$ to the NLMS algorithm. The clean reconstructed PPG signals with minimized MA were approximated by the error outputs $e[n]$ of the NLMS filter (see equations 4.1 and 4.2). In order to make real-time implementation possible, the filter was chosen to be causal. The filter length was 70 samples and an adaptation coefficient μ of 0.1 was used. Multiple reasons motivated the choice of performing ANC simultaneously on all possible ACC-PPG pairs. Indeed, it was hypothesized that the best PPG-ACC combination was time-varying and subject dependent, depending on both the sensors position and on the types of movements performed. In addition, the frequency estimation step of the proposed processing scheme is achieved by means of a multiple input adaptive frequency tracking algorithm (see Section 4.2.3). In this algorithm, the contribution of each input signal to the frequency estimate is weighted according to its relative quality. Therefore, for each time index, the best reconstructed PPG signals are automatically selected by the algorithm.

4.2.3 Instantaneous heart rate estimation

An adaptive band-pass filter was used to estimate the instantaneous frequency of the reconstructed PPG signals. More specifically, the multi-input SFT-W and the SFT-W-SPI algorithms described in Sections 2.3.2 and 2.3.3 respectively were used. The original version of the algorithm was first used and then compared with the extension including the SPI-based quality measure, which allows to adapt the filter bandwidth and convergence speed according to the local signal quality. Since these algorithms operate in the complex domain and the signals of interest were real-valued, the Hilbert transform was used to obtain the analytic representation of the latter. More specifically, for each sample, the Hilbert transform was computed on a centered window of 21 samples. For every record, the initial central frequency of the adaptive band-pass filter was computed from the maximum of the power spectral density of the first reconstructed PPG signal during the first 8 s window. For the version with adaptive δ and β parameters (SFT-W-SPI), the SPI was estimated in the time domain, using equation 2.39 (see Chapter 2). The length of the sliding window for the calculation of the SPI was two seconds. Figure 4.4 shows an example of SPI for a PPG signal. It can be noted that the SPI decreases when the quality of the PPG waveform decreases. In addition to the SPI-based mechanism, as it was empirically observed that sudden increases in signal amplitude were often related to artifacts, an additional mechanism was developed to block filter adaptation ($\delta = 1$) in such situations. For this purpose, a parameter (p_m) reflecting local amplitude increases using a 2-second sliding window was defined. For each window and for each signal indexed by m , the difference between the maximum and the minimum value was first computed. Then, p_m was computed as the ratio of the current maximum difference to the maximum difference 20-sample (0.57 s) before. Finally, δ was set to one for the next 20 samples when p_m , averaged for the available signals, was higher than a threshold empirically set at 1.7.

Figure 4.5 summarizes the different steps of the proposed HR estimation scheme.

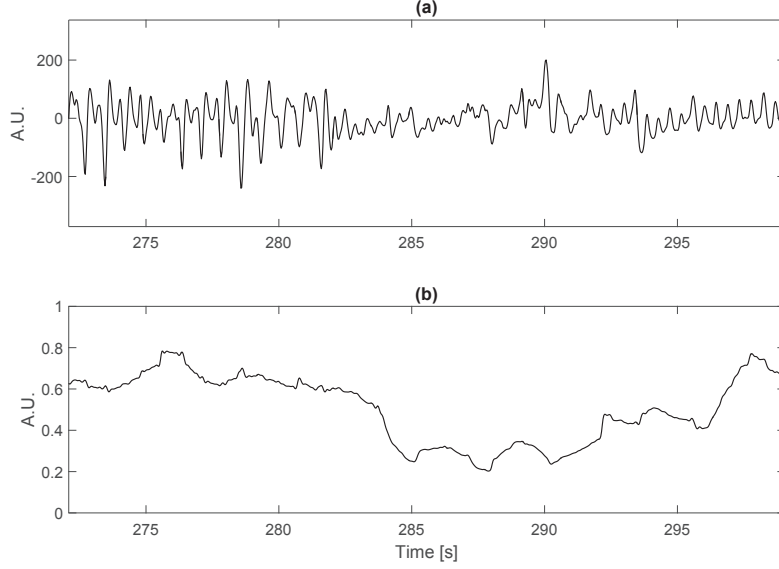


Figure 4.4: Example of SPI for a PPG signal (dataset-1, record 02). (a) PPG signal; (b) SPI.

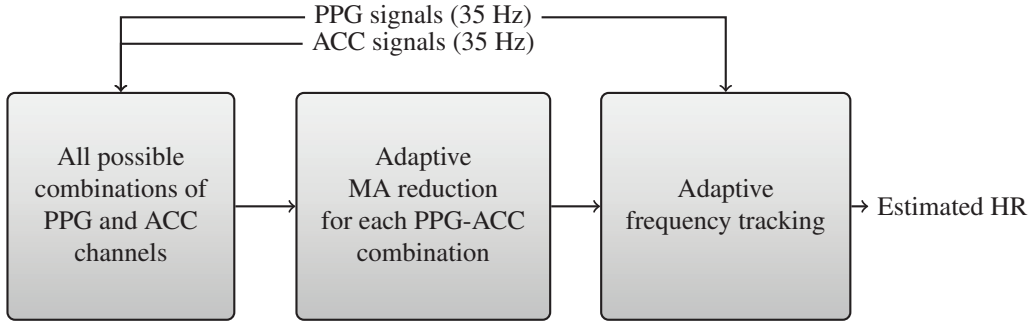


Figure 4.5: Flow diagram of the proposed HR estimation scheme.

In order to demonstrate the importance of each step of the proposed framework, HR was estimated for different configurations. More specifically, HR was estimated: with or without adaptive MA reduction, with fixed β and δ parameters (0.98) and with adaptive parameters, for all signals together and for each signal separately.

4.2.4 Performance measurement

In order to make possible the comparison with the ground-truth HR (BPM_{ref}), the obtained HR estimates were averaged on 8-second windows (BPM_{est}). Different measures were used to assess the performance of our HR estimation scheme. The average absolute error:

$$Error1 = \frac{1}{N} \sum_{i=1}^N |BPM_{est}(i) - BPM_{ref}(i)| \quad (4.3)$$

with N , the total number of time windows (see section 2.1). And the average absolute error percentage:

$$Error2 = \frac{1}{N} \sum_{i=1}^N \frac{|BPM_{est}(i) - BPM_{ref}(i)|}{BPM_{ref}(i)} \quad (4.4)$$

In order to have an additional measure of agreement between true and estimated HR values, Bland-Altman analysis was also performed. In Bland-Altman plots, the average between the ground-truth and the HR estimates is displayed on the x-axis, while the difference between these two measures is shown on the y-axis. The limits of agreement (LOA) $[\mu - 1.96\sigma, \mu + 1.96\sigma]$ to encompass 95% of the differences were calculated.

In addition to the aforementioned performance metrics, the estimation delay was also estimated. The instantaneous true HR was derived from the ECG signals (available only for the dataset-1). For this purpose, a simple local maxima detection algorithm was first used to detect the R-waves on the ECG signals. The resulting inter-beats intervals were uniformly re-sampled at 4 Hz and the instantaneous true HR was derived from these re-sampled intervals. The estimated HR was also down-sampled to 4 Hz. Then, the delay between the instantaneous true HR and the instantaneous estimated HR (using PPG signals) was computed from their cross-covariance on 30-second sliding window (28-second overlap). These delay estimates were finally averaged over all windows and all subjects.

4.3 Results

4.3.1 Performance of the presented scheme

Tables 4.2 and 4.3 present the results for the error-1 and error-2 measures, respectively, for the four different following configurations [(a)-(d)] of the proposed HR estimation scheme:

- (a) No MA reduction, HR estimation using the two raw PPG signals, adaptive β and δ .
- (b) Adaptive MA reduction (NLMS), HR estimation using all inputs (raw and reconstructed PPG signals), fixed β and δ (0.98).
- (c) Adaptive MA reduction (NLMS), HR estimation using the input leading to the smallest error, adaptive β and δ . Numbers in brackets indicate the index of the input waveforms that led to the lowest error (see table 4.4).
- (d) Adaptive MA reduction (NLMS), HR estimation using all inputs, adaptive β and δ .

Using the complete processing scheme (d), for the dataset-1, an overall average absolute error of 1.40 ± 0.60 beats-per-minute (bpm) and an overall absolute error percentage of 1.16% were obtained. Regarding the dataset-2, an overall average absolute error of 4.28 ± 3.16 bpm and an overall absolute error percentage of 4.28% were achieved. Wilcoxon signed-rank tests were performed to have some statistical comparisons between the different configurations of the proposed HR estimation scheme [(a)-(d)]. The statistical level was set to p -value < 0.01 . The error-1 was significantly reduced from configuration (a) to configuration (b), confirming the usefulness of adaptive MA reduction (p -value < 0.01). The error-1 was also significantly reduced by the use of adaptive β and δ parameters (p -value < 0.01) (comparison between (b) and (d)). The difference between (c) and (d) was not significant. Figure 4.6 shows the Bland-Altman analysis for the two datasets. LOA of $[-4.71, 4.67]$ for dataset-1 and $[-16.30, 14.35]$ for dataset-2 were obtained.

Table 4.2: Error-1 results for the four different configurations (a-d) of the presented HR estimation scheme. (a): no MA reduction, adaptive β and δ ; (b) NLMS, HR estimation using all channels, fixed β and δ (0.98); (c) NLMS, HR estimation from the best channel only, adaptive β and δ ; (d) NLMS, HR estimation using all channels, adaptive β and δ

	Record	(a)	(b)	(c)	(d)
dataset-1	1	18.28	6.95	2.05 (2)	1.75
	2	19.13	1.91	1.97 (3)	1.94
	3	1.73	1.39	1.25 (3)	1.17
	4	1.77	2.01	1.67 (3)	1.67
	5	0.91	1.07	0.99 (3)	0.95
	6	1.62	1.36	1.33 (7)	1.22
	7	1.02	1.15	0.75 (7)	0.91
	8	1.44	1.57	1.02 (3)	1.17
	9	0.92	1.19	0.85 (3)	0.87
	10	6.81	3.02	3.98 (3)	2.95
	11	8.28	1.41	1.02 (2)	1.15
	12	9.02	1.26	1.06 (4)	1.00
	Av. \pm std	5.91\pm6.68	2.03\pm1.64	1.49\pm0.89	1.40\pm0.60
dataset-2	13	12.68	12.69	10.29 (3)	12.12
	14	8.42	4.59	2.74 (1)	4.02
	15	10.44	3.70	4.85 (5)	2.52
	16	8.76	5.28	4.59 (4)	5.64
	17	19.13	3.79	2.39 (4)	3.31
	18	8.33	4.13	2.40 (3)	3.39
	19	6.08	3.36	3.55 (1)	3.45
	20	12.35	6.70	4.17 (2)	5.86
	21	2.21	1.88	1.64 (1)	1.56
	22	1.09	1.01	0.87 (3)	0.95
	Av. \pm std	8.95\pm5.24	4.71\pm3.23	3.75\pm2.63	4.28\pm3.16
total	Av. \pm std	7.29\pm6.13	3.25\pm2.78	2.52\pm2.17	2.71\pm2.58

Table 4.3: Error-2 results for the four different configurations (a-d) of the presented HR estimation scheme. (a): no MA reduction, adaptive β and δ ; (b) NLMS, HR estimation using all channels, fixed β and δ (0.98); (c) NLMS, HR estimation from the best channel only, adaptive β and δ ; (d) NLMS, HR estimation using all channels, adaptive β and δ

	Record	(a)	(b)	(c)	(d)
dataset-1	1	14.51%	6.20%	1.76% (2)	1.59%
	2	16.21%	1.87%	1.92% (3)	1.99%
	3	1.50%	1.20%	1.06% (3)	1.02%
	4	1.57%	1.81%	1.52% (3)	1.51%
	5	0.71%	0.82%	0.78% (3)	0.75%
	6	1.29%	1.16%	1.09% (7)	1.05%
	7	0.80%	0.89%	0.60% (7)	0.72%
	8	1.24%	1.36%	0.92% (3)	1.04%
	9	0.79%	1.03%	0.76% (3)	0.76%
	10	4.31%	1.98%	2.65% (3)	1.93%
	11	5.65%	0.98%	0.67% (2)	0.79%
	12	6.33%	0.97%	0.82% (4)	0.79%
	Av.	4.58%	1.69%	1.21%	1.16%
dataset-2	13	16.77%	16.94%	13.83% (3)	16.13%
	14	10.60%	5.96%	3.55% (1)	5.28%
	15	10.55%	3.13%	4.06% (5)	2.10%
	16	5.49%	3.31%	3.16% (4)	3.52%
	17	15.97%	3.19%	2.01% (4)	2.81%
	18	6.04%	3.05%	1.77% (3)	2.51%
	19	6.52%	3.99%	4.17% (1)	4.11%
	20	8.19%	4.55%	2.91% (2)	3.99%
	21	1.72%	1.47%	1.27% (1)	1.21%
	22	1.27%	1.19%	1.01% (3)	1.11%
	Av.	8.31%	4.68%	3.77%	4.28%
total	Av.	6.27%	3.05%	2.38%	2.58%

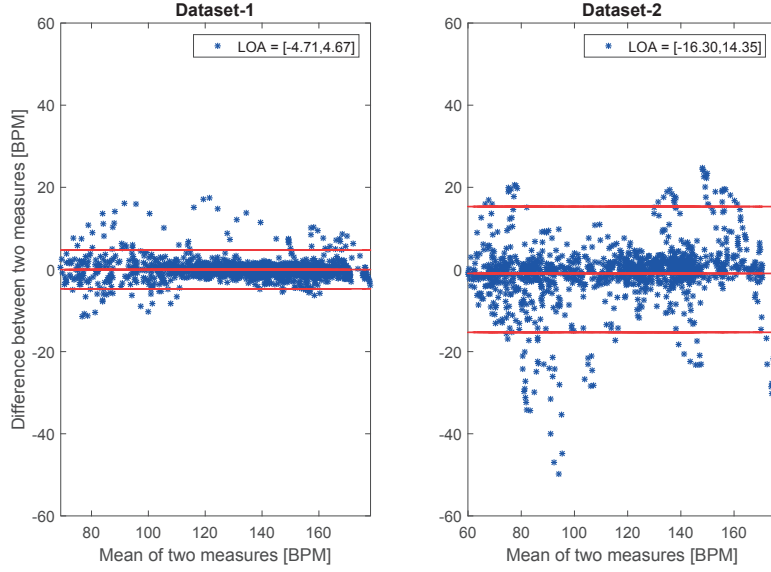


Figure 4.6: Bland-Altman plots for dataset-1 (a) and dataset-2 (b) and for the configuration (d) of the proposed HR estimation scheme. The x-axis corresponds to the mean between BPM_{ref} and BPM_{est} and the y-axis corresponds to the difference ($BPM_{ref} - BPM_{est}$).

4.3.2 Effect of adaptive MA reduction

In order to highlight the importance of the adaptive MA reduction step, the error was also reported for a configuration in which HR was estimated only from the two re-sampled raw PPG signals (Tables 4.2 and 4.3 (a)). In addition, Figure 4.7 shows an example in which the PPG signal is strongly affected by MA. The original PPG signal, as well as one of the reconstructed PPG signals are displayed on top of the ECG. The power spectral density of the two signals for the corresponding data segment is also displayed. Figure 4.8 shows the HR estimate, as well as the ground-truth, for a record of dataset-1, with and without adaptive MA reduction.

4.3.3 Combining inputs

The columns (c) of Tables 4.2 and 4.3 emphasize the interest of performing adaptive MA reduction for every possible PPG-ACC combination and using all the reconstructed PPG signals as inputs for adaptive frequency tracking. For each subject, the HR was estimated on the different reconstructed and original PPG channels separately and the one resulting in the lowest average absolute error is displayed. The numbers in brackets correspond to the following channel combinations: 1: ACC(1) & PPG(1), 2: ACC(2) & PPG(1), 3: ACC(3) & PPG(1), 4: ACC(1) & PPG(2), 5: ACC(2) & PPG(2), 6: ACC(3) & PPG(2), 7: PPG(1) only, 8: PPG(2) only. Table 4.4 summarizes the number of times each reconstructed/raw PPG signal led to the best HR estimation in terms of average absolute error.

4.3.4 Effect of adaptive β and δ

In order to highlight the effect of adapting δ and β , the error measures have been computed also for fixed δ and β (Tables 4.2 and 4.3 (b)). Figure 4.9 illustrates the evolution of error-1 as a function of δ and β for three different records of the database.

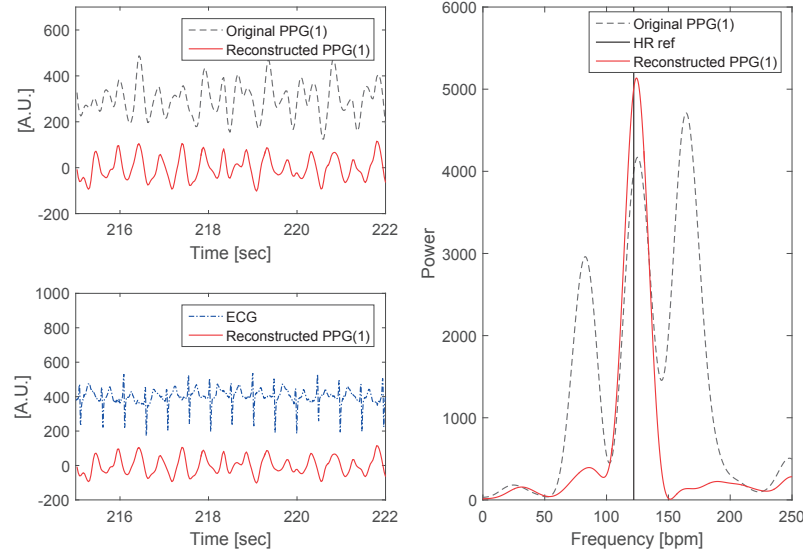


Figure 4.7: Effect of adaptive MA reduction on a data segment (record 02, dataset-1). (a) Original PPG(1) and reconstructed PPG(1); (b) ECG and reconstructed PPG(1); (c) Power spectral density of the original PPG(1) and reconstructed PPG(1) signals.

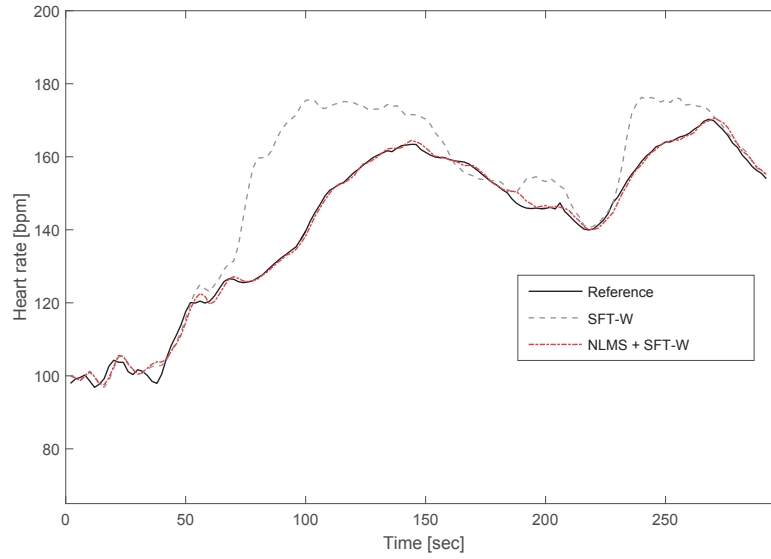


Figure 4.8: HR estimation with and without adaptive MA reduction (record-12).

4.3.5 Difference between the different types of physical exercises

Figure 4.10 compares the performance of our method, in terms of average absolute error, for the different types of physical exercise that were performed by the subjects. These error values correspond to the final HR estimation scheme (d).

Table 4.4: Number of times each channel combination led to best heart rate estimation.

Combination number	Waveforms	Number of times it gave the best estimation
1	ACC(1) & PPG(1)	3
2	ACC(2) & PPG(1)	3
3	ACC(3) & PPG(1)	10
4	ACC(1) & PPG(2)	3
5	ACC(2) & PPG(2)	1
6	ACC(3) & PPG(2)	0
7	PPG(1)	2
8	PPG(2)	0

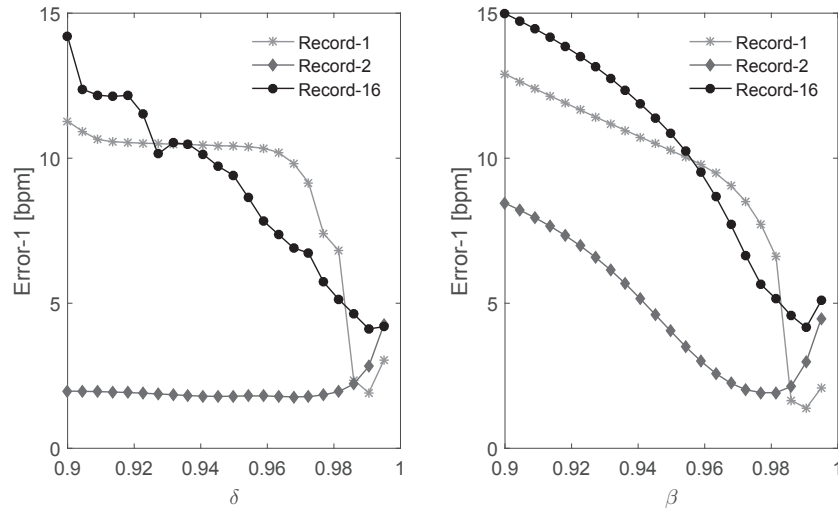
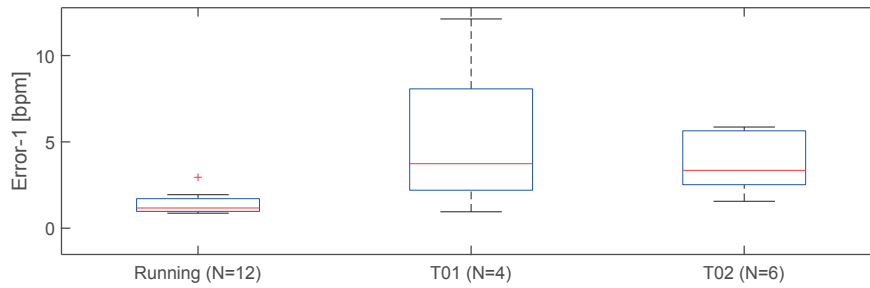
Figure 4.9: Left side: Relationships between the error-1 and δ for three records (β fixed at 0.98); right side: Relationships between error-1 and β for the same records (δ fixed at 0.98).

Figure 4.10: Error-1 results for the different types of physical activities performed by the subjects.

4.3.6 Estimation delay

An average estimation delay of 0.93 ± 0.30 seconds was found between the true and the estimated HR (final estimation scheme (d)). It should be noted that this delay is variable and depends on different parameters, such as the value of parameter δ in the SFT-W algorithms and the nature of the frequency changes. For example, a larger delay was observed for epochs containing very abrupt HR changes. Importantly, as the cross-correlation was computed between HR estimated from ECG and HR estimated from PPG. A fraction of this estimated delay corresponds to the pulse transit time, which varies according to different factors such as age, HR, height, and systolic blood pressure [53].

4.3.7 Comparison with the state-of-the-art

Results from some of the studies mentioned in the introduction are reported in Table 4.5. Wilcoxon signed rank tests were performed to compare the results. Regarding dataset-1, the error values obtained using our scheme were significantly lower than that of TROIKA [43] (p -value < 0.01), but significantly larger than that of MISPT [45] (p -value < 0.05). There was no significant difference between our scheme and the JOSS [44] and SPECTRAP [47] ones. Regarding the second dataset, our error values were significantly larger than that of SPECTRAP (p -value < 0.01) [47].

Table 4.5: Comparison with existing studies (error-1).

	Record	TROIKA [43]	JOSS [44]	SPECTRAP [47]	MISPT [45]	Our
dataset-1	1	2.29	1.33	1.18	1.58	1.75
	2	2.19	1.75	2.42	1.80	1.94
	3	2.00	1.47	0.86	0.58	1.17
	4	2.15	1.48	1.38	0.99	1.67
	5	2.01	0.69	0.92	0.74	0.95
	6	2.76	1.32	1.37	0.93	1.22
	7	1.67	0.71	1.53	0.73	0.91
	8	1.93	0.56	0.64	0.45	1.17
	9	1.86	0.49	0.6	0.41	0.87
	10	4.70	3.81	3.65	3.60	2.95
	11	1.72	0.78	0.92	0.88	1.15
	12	2.84	1.04	1.25	0.68	1.00
	Av. \pm std	2.34 \pm 0.83	1.29 \pm 0.90	1.39 \pm 0.86	1.11 \pm 0.89	1.40 \pm 0.60
dataset-2	13	-	-	4.89	-	12.12
	14	-	-	1.58	-	4.02
	15	-	-	1.83	-	2.52
	16	-	-	3.05	-	5.64
	17	-	-	1.62	-	3.31
	18	-	-	1.24	-	3.39
	19	-	-	2.04	-	3.45
	20	-	-	2.49	-	5.86
	21	-	-	1.16	-	1.56
	22	-	-	0.66	-	0.95
	Av. \pm std	-	-	2.06 \pm 1.21	-	4.28 \pm 3.16

4.4 Discussion

The overall error-1 of 1.40 ± 0.60 [bpm] (error-2 of 1.16 %) for dataset-1 and 4.28 ± 3.16 [bpm] (error-2 of 4.28%) for dataset-2 are encouraging. Columns (a) of Tables 4.2 and 4.2 indicate that PPG signals alone were not sufficient to compute accurate HR estimates and the use of ACC signals was necessary. It should be observed that MA reduction was more efficient for the first dataset, for which an error-1 reduction of 76% was achieved when NLMS was used prior to adaptive frequency tracking. For the second dataset, the use of NLMS resulted in an error-1 reduction of 52%. An example of effective MA reduction is shown in Figure 4.7, where the original PPG signal is affected by MA and exhibits supplementary pulses whereas the reconstructed PPG signal is well synchronized with the ECG. This was confirmed by the power spectral density of these two signals, in which only one clear peak at the true HR was visible after adaptive MA reduction.

Regarding the contribution of each input signal (raw and reconstructed PPG signals), different conclusions can be drawn from Table 4.4. The first PPG channel (PPG(1)) led to better results than the second one (PPG(2)), despite the similarity between the two sensors (green LEDs at 609 nm, distance of two centimeters between them). A possible explanation for this difference of performance is that the position of the first LED (PPG(1)) was more appropriate than that of the second one (better skin contact, more blood vessels, etc.). It should also be noted that the input signal resulting in the smallest error varies for the different recordings and can not be known in advance. Therefore, the possibility of easily combining all the inputs in adaptive frequency tracking is of great interest.

The possible benefit of adaptive δ and β parameters was also investigated in this study. The HR estimation was significantly improved by adaptive δ and β parameters. More precisely, average error-1 reductions of 31% and 9% were achieved for dataset-1 and dataset-2, respectively (Table 4.2). In addition to this error reduction, the interest of using adaptive β and δ is illustrated in Figure 4.9, where it can be noticed that the optimal value for β and δ , in terms of error-1, varies for the different records. In addition, very slight changes of β and δ can lead to important change of performance (for example record-1). Regarding the underlying adaptive mechanism, SPI seems to be a good candidate, as it is representative of the PPG signal quality (see example in Figure 4.4). Indeed, a good quality PPG signal presents a main oscillation at HR frequency, while a PPG signal corrupted by MA shows oscillations at various frequencies.

It should be noticed that better results were achieved by our method for the first dataset, as shown by the Bland-Altman plots (Figure 4.6). This may be explained by the fact that MA induced by running are more or less periodic and stationary and thus easier to attenuate. Very abrupt transitions can be observed in the ACC signals for the second dataset and this may be the reason why the NLMS was less efficient for dataset-2 and larger error values were obtained. Surprisingly, better results were obtained for exercise type T02 (Figure 4.10), which was more intensive than exercise type T01. This dataset allowed us to highlight the limitations of the presented scheme. For example, we can notice that HR estimation failed for the first recording of dataset-2 (Table 4.2).

Our approach provides results that are comparable to the state-of-the-art approaches for dataset-1 (Table 4.5). However, our scheme is less accurate than the SPECTRAP one [47] for the dataset-2. Surprisingly, the authors of the other studies did not report the performance of their algorithms on dataset-2, which was in fact more challenging than dataset-1.

In addition to the aforementioned results, our method has the advantage of being almost real-time, while the studies mentioned have a 4-second delay. This real-time aspect is frequently mentioned as a requirement in the development of smart wearable systems. Indeed, intelligent medical monitoring devices are often described as devices capable of providing real-time processing and feedback to medical staff, patients, athletes, and healthy subjects [54].

Another strength of the presented HR estimation scheme is the very small number of parameters

to tune, rendering the methods easily applicable to different data and different experimental situations.

Finally, it should be mentioned that our method does not crucially depend on the initialization step as it is the case in other studies [43, 44].

4.5 Conclusion

In this chapter, we propose a new approach to estimate HR using PPG and ACC signals for subjects performing different kind of physical activities. Our results are in the same range as the ones resulting from the state-of-the-art algorithms in this field. However, unlike the previously proposed Fourier-based approaches, our scheme provides continuous heart rate estimates almost in real-time, which is of great interest for real-time applications. In addition, all the signals are automatically combined in an optimal way and thus, no a priori knowledge about the contribution of the different waveforms is required. Finally, only a very few number of parameters have to be tuned, making this approach suitable and flexible for different experimental situations. Further work should concentrate on improving the tracking during physical exercises similar to those of dataset-2. In addition, the inclusion of a confidence index on the estimated HR values could be valuable.

Atrial Fibrillation Detection

5

5.1 Introduction

The present chapter aims at evaluating the potential of different PPG-derived measures to discriminate between different types of rhythm, namely atrial fibrillation (AF), ventricular arrhythmia (VA) and sinus rhythm (SR), in the perspective to perform atrial fibrillation detection using the PPG signal. After defining atrial fibrillation and reviewing the literature about atrial fibrillation detection using PPG, novel PPG-wave measures reflecting signal organization/complexity are presented and their performance in detecting atrial fibrillation is evaluated. These measures are also compared with more conventional measures based on heart rate variability.

5.1.1 Definition of atrial fibrillation

Atrial fibrillation (AF) is a supraventricular tachyarrhythmia characterized by uncoordinated atrial activation, which results in the deterioration of atrial mechanical functions. As illustrated in Figure 5.1 AF can be identified on the ECG. It can be seen that the P-waves are replaced by rapid oscillations, referred to as fibrillatory waves. These fibrillatory waves have variable amplitude, morphology and period. Their frequency is usually between 300 and 600 bpm [55].

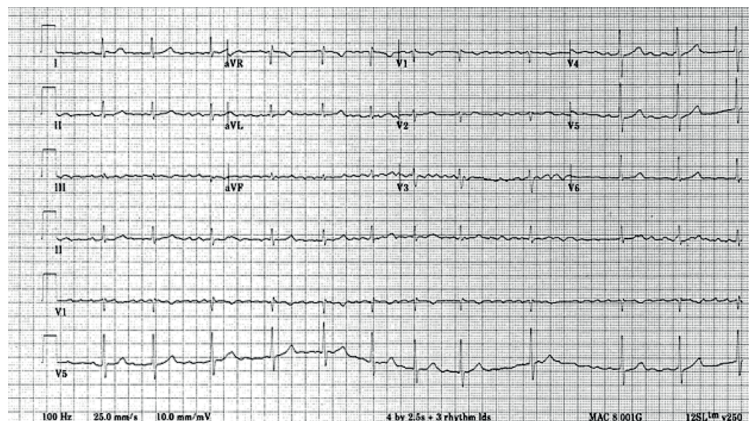


Figure 5.1: Example of ECG showing AF. Figure from [55].

AF is the most common cardiac arrhythmia, affecting 1-2% of the general population. Among several AF-related cardiovascular outcomes, AF is known to increase stroke risk [56], and thus should be diagnosed early. AF may present as a unique episode or be recurrent. Different classifications of AF have been established to facilitate the management of patients. Depending on episodes duration, recurrent AF is further classified as: paroxysmal if it terminates spontaneously within 7 days, persistent if sustained beyond 7 days (<7 days if cardioversion is performed) or longstanding persistent in case of a continuous episode lasting more than one year. Finally AF is defined as permanent if cardioversion has failed or was not attempted [56]. AF can be symptomatic or asymptomatic. Patients often experience a sensation of palpitations and other symptoms such as chest pain, dyspnea or dizziness have also been reported.

The paroxysmal nature of AF in its initial course renders its diagnosis challenging and requires long-term monitoring to capture AF episodes. The gold standard diagnostic tools are Holter and implantable cardiac monitors. The latter are widely accepted to be the most reliable monitoring method to assess AF therapies [57]. However, implantable cardiac monitors require an invasive procedure which limits their use on a large population. On the other hand, efforts in recent years have focused on providing low cost, portable and minimally invasive diagnostic tools that are suitable for diseases screening on a large scale.

5.1.2 Literature review on the use of photoplethysmographic sensors to detect atrial fibrillation

The principle of PPG was introduced in Chapter 3. Because they can be comfortably worn during day-life activities, a lot of wearable PPG-based heart rate monitors have recently emerged on the market. For example, we have seen in the previous chapter that PPG-wrist type devices can be used to estimate HR during sport sessions. However, despite this attractiveness for PPG devices, their potential as diagnostic tools to detect cardiac arrhythmias is still partly unexplored. More specifically, the number of studies reporting the performance of AF detection using PPG is still quite limited. In [58, 59], the PPG signals provided by an iPhone 4S were used to detect AF. The interbeat intervals (IBI) were first computed from the PPG waveforms and some statistical measures were derived from the IBI series, including the normalized root mean square of successive difference of IBI (nRMSSD), Shannon entropy and a parameter derived from a Poincaré plot. In another study, AF detection using a wrist-type PPG sensor was investigated by Bonomi *et al.* [60]. More specifically, 16 patients were monitored for 24 hours with a standard Holter monitor and a wrist-wearable PPG device. AF detection was performed using the IBI series derived from the PPG signals. A first-order Markov model was used successfully to compute the probability of AF based on the pattern of IBI series. Corino *et al.* investigated the use of the blood volume pulse signal to discriminate between AF, other arrhythmias and SR [61]. In this study 24 features were computed. They can be categorized in three different classes: spectral analysis, variability and irregularity analysis and shape analysis. The two first classes are derived from the IBI (the inter-systolic and the inter-diastolic IBIs were studied). Highest variability parameters were obtained for AF, and the most relevant features were pNN40 and pNN70¹. Using these two features, specificities of 0.928/0.963/0.768 and sensitivities of 0.773/0.754/0.758 were obtained for AF, SR and other arrhythmias, respectively. In another study by Conroy *et al.*, four HRV parameters were derived from the inter-systolic intervals computed on the PPG signal [62]. In this study, the PPG signal was recorded using an earlobe PPG sensor. Sensitivity and specificity of 90.9% were obtained for the pNN35 parameter.

As for AF, a PPG-based monitoring device for ventricular arrhythmia (VA) detection would present a substantial clinical interest. There are relatively fewer research works on VA detec-

1. pNNx: The percentage of successive normal cardiac interbeat intervals greater than x msec.

tion. PPG-based detection of heart rate turbulence [63], that is, the short-term oscillation in heart rate induced by a PVC has been proposed. Turbulence onset and slope, as well as turbulence-induced wave shape changes were used. Another work proposed a PVC-detection scheme based on features characterizing PPG pulse power and peak-to-peak intervals [64]. To the best of our knowledge, only one publication dealt with the problem of PPG-based classification of VA versus sinus rhythm (SR) and supraventricular (i.e., mainly atrial) premature contractions [65].

The present study aims at evaluating the potential of a PPG wrist-type device to discriminate between different types of rhythm, namely AF, VA and SR. In preliminary studies, we showed that: 1) a few PPG-IBI features can relatively well discriminate between AF and SR [66] and 2) the organization level of PPG signal, which can be assessed through the adaptive organization index, is significantly lower during AF [67].

In the present study, we introduce a set of features consisting of complexity/organization measures, completely innovative in the context of PPG processing, and we evaluate their ability to discriminate between AF and the other rhythms. We show that, besides being as discriminative as the PPG-IBI features to classify AF and SR episodes, the proposed features lead to better performance for the classification of AF and VA episodes.

5.2 Methods

5.2.1 Database

The database consists of 17 consecutive patients (12 men, 5 women, age: 57 ± 13 years old) referred for catheter ablation of cardiac arrhythmia (Arrhythmia Unit, Heart and Vascular Department, Lausanne University Hospital, Switzerland). For each patient, the following waveforms were continuously recorded during the ablation procedure: ECG, PPG and three-axis accelerometer signals (ACC). More specifically, the 12-lead ECG was acquired at 2 kHz (Sensis, Siemens) and the PPG signal was recorded using a wrist-type device developed by the Swiss Center for Electronics and Microtechnology (CSEM). This PPG device is composed of a LED operating at near-infrared light in reflection mode and an embedded accelerometer. The resulting ACC and PPG signals were sampled at 21.33 Hz.

5.2.2 Annotation of ECGs

In order to define a ground-truth for the different types of rhythm, ECG waveforms were annotated by a team of local experts for the segments belonging to one of these three classes: SR, AF and VA. The VA class consisted of ventricular tachycardia, premature ventricular contractions and bigeminy episodes. Table 5.1 summarizes the number of annotated 10-second epochs for each class. A total of 2166 labeled epochs was obtained for the three classes (SR, AF and VA).

5.2.3 Pre-processing

Alignment of PPG and ECG signals

ECG and PPG waveforms were temporally aligned using the IBI series extracted from PPG and ECG [66]. Regarding the ECGs, the detection of R-waves was performed directly by the recording system (Sensis, Siemens) and the RR intervals were computed from the differences between consecutive R-waves. Regarding the PPG signals, the systolic downstrokes of each pulse were identified as the zeros of the PPG signal second derivative. A finite impulse response

Table 5.1: Number of epochs of SR, AF and VA, for each patient.

Patient	SR	AF	VA	Total
1	107	0	0	107
2	26	0	74	100
3	0	0	25	25
4	63	0	0	63
5	82	0	170	252
6	0	100	0	100
7	0	512	0	512
8	20	0	0	20
9	0	22	0	22
10	13	0	110	123
11	0	0	6	6
12	0	323	0	323
13	12	0	6	18
14	0	156	0	156
15	0	166	4	170
16	0	91	0	91
17	58	0	20	78
Total	381	1370	415	2166

filter with the following coefficients was first applied to the raw PPG signal:

$$dd_n = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{-2}{3}, \frac{-2}{3}, \frac{-2}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \odot \text{hanning}(9) \quad (5.1)$$

Where \odot is the element-wise multiplication and $\text{hanning}(9)$ is a Hanning window of length 9. Then, pairs of samples crossing the zero-value were identified from the second order derivative and were recognized as systolic downstrokes. In order to increase the temporal resolution, a linear interpolation was applied for each pair of samples.

PPG-based IBI were computed from the time difference between consecutive events. For each record, a beat-to-beat alignment of the ECG- and PPG-based IBI series was performed. A dynamic time warping algorithm was used for this purpose [68]. After the alignment of the ECG- and PPG-based IBI series, the relation between the time-bases of the two recording systems was assumed to be modeled by the following linear equation:

$$\widehat{t_{ECG}} = \alpha \cdot t_{optical} + \beta \quad (5.2)$$

where $\widehat{t_{ECG}}$ is the estimated time in the ECG system time base, $t_{optical}$ is the estimated time in the PPG system time base, α is the relative drift of the PPG system clock relatively to the ECG system clock, and β is the time offset of the two recordings. The parameters α and β were estimated using classical regression. Given the noisy nature of the PPG-based IBI series, removal of outliers prior to the estimation of the model parameters was necessary. The RANSAC algorithm was used for this purpose, with the residual threshold set to one second [69]. Finally, equation (5.2) was used to transform the time-base of the PPG signals to match the time-base of the ECG signals².

Removal of motion artifacts

As discussed in Chapters 3 and 4, PPG waveforms are very sensitive to motion artifacts and different approaches have been proposed to perform adaptive cancellation [42]. In this study, we

2. The alignment of PPG and ECG signals was performed by Mathieu Lemay and Philippe Renevey, CSEM.

decided to discard all the PPG segments contaminated with motion artifacts in order to assess the efficiency of our methods to detect arrhythmias in optimal conditions. The ACC signals were used in this regard. More specifically, when the local standard deviation of the ACC signals was larger than a fixed empirically selected threshold, the corresponding PPG regions were not considered for further analysis. Figure 5.2 illustrates this detection of motion artifacts.

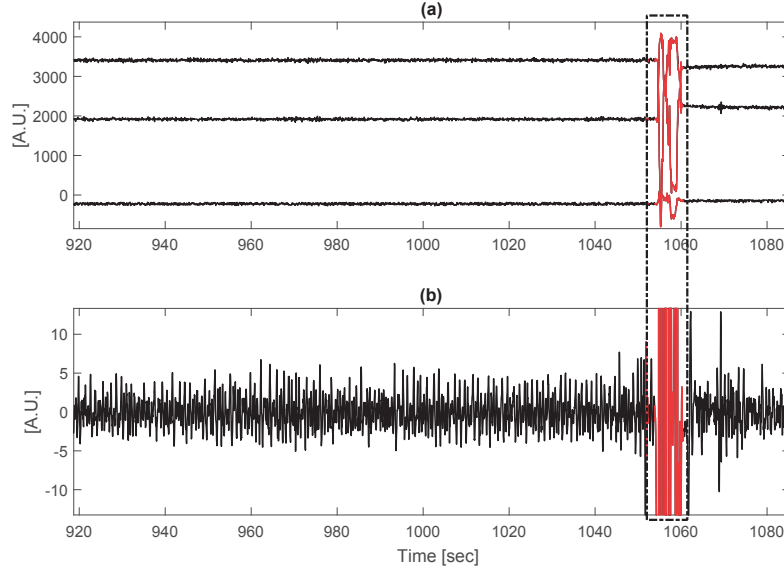


Figure 5.2: Removal of PPG segments contaminated by motion artifacts. (a) ACC signals; (b) Pre-processed PPG signal. Example of a segment automatically detected as a motion artifact.

5.2.4 Computation of the features

This subsection describes the features that were computed on the 10-second epochs of PPG signal (no overlap) corresponding to the annotated ECG sections. Two different types of features were computed; the PPG-wave features and the PPG-IBI features. Based on the simple observation that the PPG waveform is a lot more irregular in terms of morphology and amplitude during AF episodes, we hypothesized that AF leads to changes in hemodynamics such as reduced stroke volume (followed by a reduction in systemic blood pressure) and a pooling of venous blood, leading to such pathological waveform morphology. We therefore investigated PPG-wave features able to reflect such changes. The PPG-IBI features, on the other hand, were computed from the PPG-IBI series.

PPG-wave features

- **Adaptive organization index (AOI):** the organization index, introduced in [70], is defined as the ratio of the combined power of the fundamental frequency and the first harmonic of an oscillation to the total power of the signal. An adaptive method for continuous and instantaneous estimation of this index was proposed in [71]. In this study, the authors used the harmonic frequency tracker described in [21], based on time-varying single pole band-pass filters to track the instantaneous fundamental and first harmonic frequency components. The power was then computed from the low-pass version of the extracted compo-

nents squared (fundamental and first harmonic). Using the aforementioned methodology to compute the AOI, we showed in a preliminary study [67] that the AOI is representative of PPG signal organization by comparing its average value for four different types of rhythm, namely: SR, regularly paced rhythm, irregularly paced rhythm (atrial and ventricular pacing protocols with increment in pacing rate) and AF. This is shown in Figure 5.3. In the present study, the average AOI value for each 10-second epoch considered was reported as a feature.

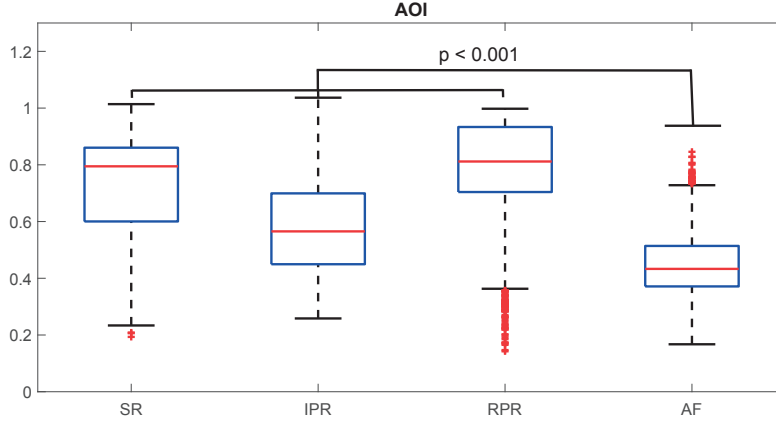


Figure 5.3: Boxplot of the obtained AOI values for the four classes, from our preliminary study [67]. SR: sinus rhythm, IPR: irregularly paced rhythm, RPR: regularly paced rhythm, AF: atrial fibrillation.

- **Variance of the slope of the phase difference (VSPD):** the phase difference between the harmonic components of a signal reflects their synchronization and, therefore, is an indication of signal organization. Using the same harmonic frequency tracker [21], the instantaneous phase difference between the fundamental and the first harmonic was computed using the Hilbert transform. The variance of the slope of the phase difference (VSPD) was calculated for each 10-second epoch. Larger VSPD values were expected during AF, reflecting a reduced synchronization between the fundamental component and its first harmonic.
- **Permutation entropy (PE):** permutation entropy is a complexity measure based on the probability distribution of the permutation motifs [72]. The permutation entropy of order n ($n \geq 2$) is defined as:

$$PE(n) = - \sum p(\pi) \log(p(\pi)) \quad (5.3)$$

With $p(\pi)$ the relative frequency of the permutation motif π :

$$p(\pi) = \frac{|\{t | t \leq T - n, (x_{t+1}, \dots, x_{t+n}) \text{ has type } \pi\}|}{T - n + 1} \quad (5.4)$$

with T the length of a signal x . PE has been used in biomedical applications, such as detection of dynamic changes in the analysis of electroencephalogram signals [73]. In the present study, the PE was computed on the PPG waveform for each 10-second epoch, using $n = 4$. Larger PE values were expected during arrhythmias.

- **Spectral entropy (SE):** spectral entropy quantifies the spectral complexity of a time series. It is defined as the Shannon entropy of the normalized power spectral density function [74]:

$$SE = \sum_f p_f \log\left(\frac{1}{p_f}\right) \quad (5.5)$$

With p_f , the power spectral density normalized with respect to the total spectral power. In this study, the power spectral density was computed using an Fourier-based method (Welch's) as recommended in [74]. Figure 5.6 illustrates the SE values obtained for a sinusoid ($\omega_0 = 0.2\pi$) with additive Gaussian white noise at different SNR values.

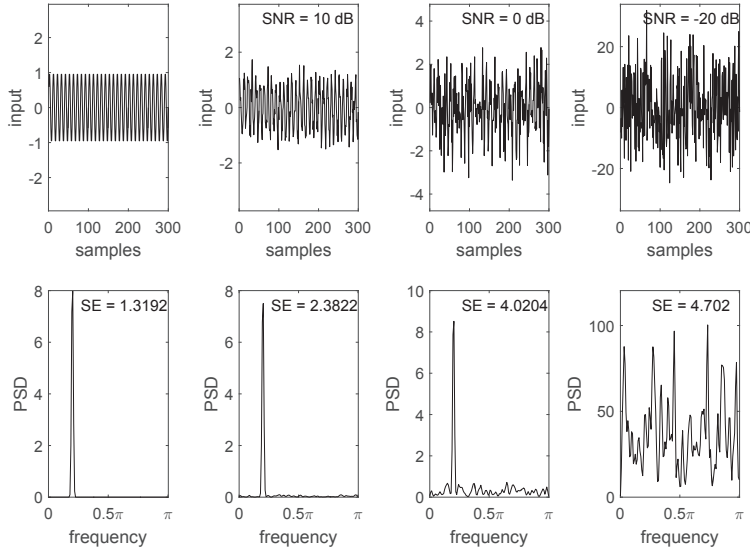


Figure 5.4: This example illustrates the SE values obtained for a sinusoid contaminated with different level of white Gaussian noise.

As for PE, this measure has mainly been used in biomedical applications involving the analysis of electroencephalogram signals [74]. In this study, we hypothesized that a higher spectral complexity would be observed during AF episodes, and thus higher SE values.

- **Fractional spectral radius (FSR):** a vector u_i is first defined as an embedding vector whose elements are m samples taken at intervals of J samples along the observed time series x :

$$u_i = (x_i, x_{i+J}, \dots, x_{i+(m-1)J})^T, i = 1 : N - (m-1) \cdot J \quad (5.6)$$

with m , the embedding dimension. One can construct an embedding matrix U from the embedding vectors, such that the lines of U are $u_1^T, u_2^T, \dots, u_{N-(m-1) \cdot J}^T$. Then, the eigenvalues σ_i^2 of $U^T U$ are computed and their spread can be quantified using the fractional spectral radius (FSR), defined as [74]:

$$FSR(k) = \frac{\sum_{i=0}^k \sigma_i^2}{\sum_{l=0}^m \sigma_l^2} \quad (5.7)$$

In the study, the value of k was empirically set to 1 and we used $J = 11$ and $m = 3$. Smaller FSR values were expected during AF, indicating a higher complexity.

- **Spectral purity index (SPI):** the SPI measure described in Section 2.4. In this study, a sliding window of length $L = 40$ samples was used to compute the instantaneous SPI. Then, for each epoch, the average SPI was reported as feature. Low SPI values were expected during AF episodes.

PPG-IBI features

The following PPG-IBI features were computed for each IBI series corresponding to a 10-second annotated PPG epoch:

- Mean: $\text{mean}(\text{IBI})$
- Standard deviation: $\text{std}(\text{IBI})$
- Median: $\text{median}(\text{IBI})$
- Interquartile range: $\text{iqr}(\text{IBI})$
- Minimum: $\text{min}(\text{IBI})$
- Maximum: $\text{max}(\text{IBI})$
- Square root of the mean squared differences of successive IBI [75]: RMSSD

5.2.5 Classification

Bagging decision trees were used to assess the discriminative performance of the aforementioned features. Three groups of features were considered: PPG-wave features, PPG-IBI features and all features together. The performance of each feature group was evaluated for the three following classification schemes: AF against SR, AF against VA and AF against (SR & VA). In order to reduce the variability of the results, a k-fold cross-validation was performed with $k = 5$.

5.2.6 Accuracy of the computed IBI series

Heartbeat detection in PPG signals may be a challenging task, especially in the presence of arrhythmias. In order to evaluate the efficiency of the heartbeat detection method used here (see Sect. 5.2.3), the detected heartbeats were compared to the reference heartbeats directly provided by the ECG recording device (Siemens, Sensis). The heartbeats detected from the PPG signal were classified into three categories: true positive (TP), false positive (FP) and false negative (FN). Then, the detection error rate (DER) was used as a performance metric:

$$\text{DER} = \frac{\text{FP} + \text{FN}}{N_{\text{peaksref}}} \quad (5.8)$$

where N_{peaksref} indicates the total number of heartbeats in the reference signal. The overall DER was estimated over all the 10-second epochs used to calculate the features. In addition the DER was reported separately for the three categories (SR, AF, VA).

5.3 Results

5.3.1 Computed features

Figure 5.5 is a representative example illustrating the steps involved in the computation of the AOI. Figure 5.6 illustrates the normalized PSD of two PPG segments during SR and AF. The resulting SE, computed from the PSD, is also displayed. Boxplots for the different features are shown in Figure 5.7. Table 5.2 reports the AUC of the ROC for each feature taken separately and for two classification schemes: AF against SR and AF against VA.

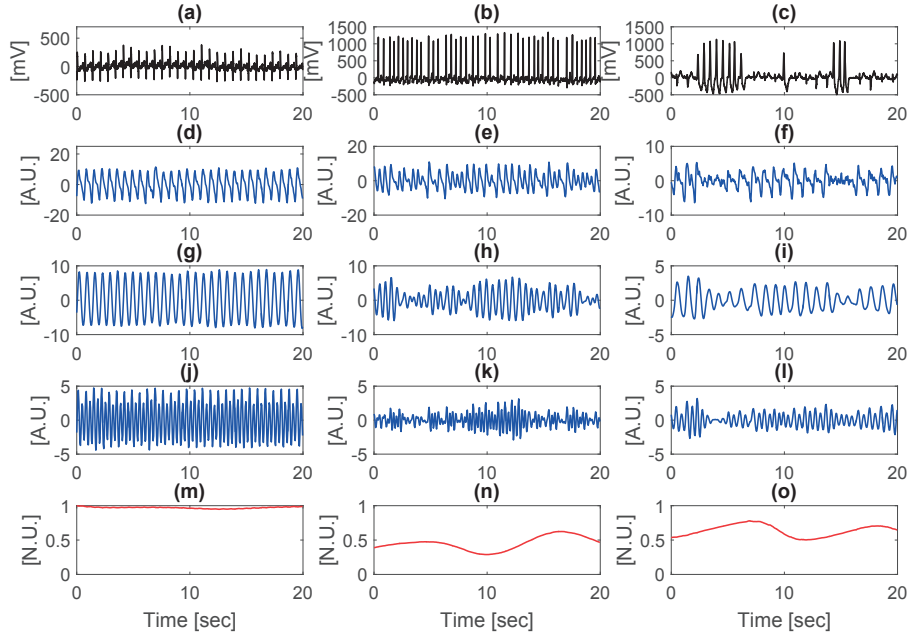


Figure 5.5: Example illustrating the computation of the AOI feature for the different types of rhythms: SR, AF and VA (from the left to the right). (a-c) ECG waveforms; (d-f) PPG waveform; (g-i) Fundamental components; (j-l) First harmonic components; (m-o) Resulting AOIs.

The results displayed in Table 5.2 indicate that some features alone could already discriminate well between SR and AF. However, it was not the case for the classification of AF and VA. Therefore we decided to investigate classification schemes involving combinations of the proposed features.

5.3.2 Classification performance:

The accuracy, specificity and sensitivity values obtained for the classification of AF and SR events and for the different groups of features are reported in Table 5.3.

The same values for the classification of AF and VA events and for the different groups of features are reported in Table 5.4.

Table 5.5 reports the same performance metrics for the classification of AF and other events and for the different groups of features.

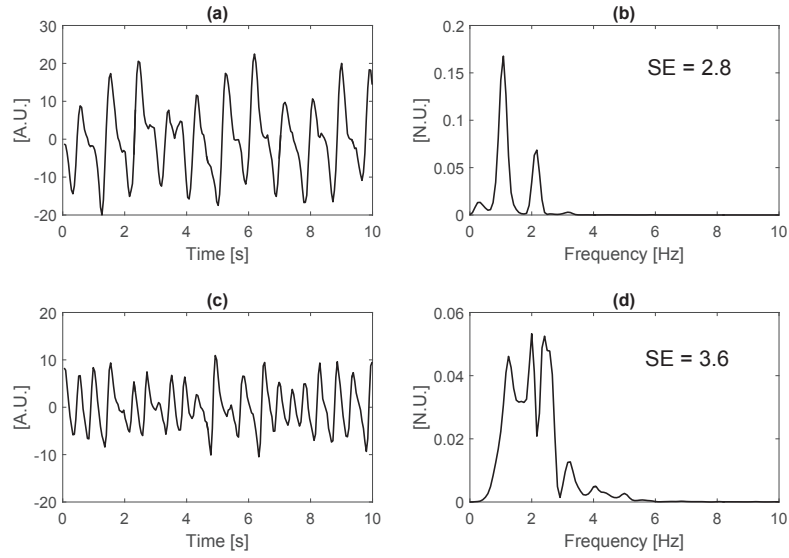


Figure 5.6: Example of PPG waveforms (a,c), their PSD (b,d) and the resulting SE values. Top row: SR; bottom row: AF.

Table 5.2: AUC values for the different features.

Feature	AUC SR-AF	AUC AF-VA
AOI	0.92	0.74
VSPD	0.93	0.70
PE	0.92	0.56
SPI	0.60	0.84
SE	0.87	0.52
FSR	0.81	0.66
Mean(IBE)	0.68	0.60
Std(IBE)	0.88	0.52
Median(IBE)	0.72	0.58
Iqr(IBE)	0.93	0.54
Min(IBE)	0.83	0.62
Max(IBE)	0.78	0.65
RMSSD	0.90	0.54

Table 5.3: Classification performance AF against SR.

	Accuracy [%]	Specificity [%]	Sensitivity [%]
PPG-wave features	97.3	90.6	99.2
PPG-IBE features	97.3	89.5	99.5
All features	98.1	92.4	99.7

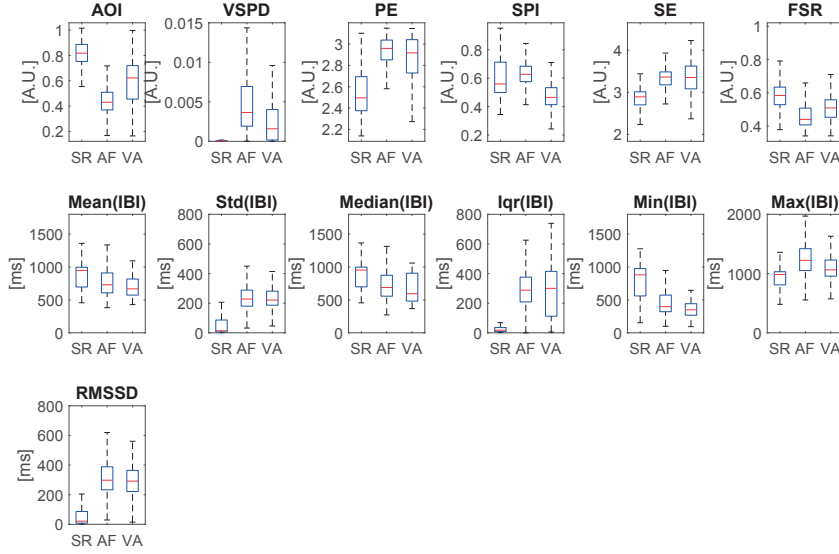


Figure 5.7: Boxplots for all features, computed from all the labeled PPG epochs after removal of segments corresponding to motion artifacts. AOI: adaptive organization index; VSPD: variance of the slope of the phase difference; PE: permutation entropy; SPI: spectral purity index; SE: spectral entropy; FSR: fractional spectral radius; Mean(IBE): mean IBE, Std(IBE): standard deviation IBE; Median(IBE): median IBE; Iqr(IBE): interquartile range IBE; Min(IBE): minimum IBE; Max(IBE): maximum IBE; RMSSD: square root of the mean squared difference if successive IBE.

Table 5.4: Classification performance AF against VA.

	Accuracy [%]	Specificity [%]	Sensitivity [%]
PPG-wave features	93.9	87.0	96.1
PPG-IBE features	88.1	63.1	95.7
All features	95.9	88.7	98.1

Table 5.5: Classification performance AF against (SR&VA).

	Accuracy [%]	Specificity [%]	Sensitivity [%]
PPG-wave features	93.8	91.5	95.1
PPG-IBE features	88.5	78.4	94.3
All features	95.0	92.8	96.2

5.3.3 Comparison with IBE-ECG-derived features

Table 5.6 compares the classification performance for the IBE features derived separately from PPG and ECG.

5.3.4 Accuracy of the computed PPG-IBE series

Table 5.7 reports the DER obtained for the different classes.

Table 5.6: Classification performance for ECG

		Accuracy [%]	Specificity [%]	Sensitivity [%]
AF/SR	PPG-IBI features	97.3	89.5	99.5
	ECG-IBI features	99.4	98.2	99.8
AF/VA	PPG-IBI features	88.1	63.1	95.7
	ECG-IBI features	95.9	89.2	97.9
AF/(SR&VA)	PPG-IBI features	88.5	78.4	94.3
	ECG-IBI features	96.4	93.7	98.0

Table 5.7: DER of the detected heartbeats in PPG signal.

All	SR	AF	VA
0.28±0.31	0.05±0.11	0.33±0.32	0.33±0.28

5.4 Discussion

The present study aimed at evaluating the potential of a set of PPG-derived measures to discriminate between AF and other types of rhythm. Such approaches could improve the screening of AF by rendering it less cumbersome, non invasive and at a lower cost.

As mentioned in the introduction, the number of studies dealing with AF detection using PPG is limited. In these studies, promising results were reported for the discrimination between AF and SR rhythms. In [60], the classification of 30-seconds segments led to an accuracy of 98%. An AUC of 0.966 was obtained in [59], and an accuracy of 96.8% was achieved in [58]. Our results reported in Table 5.3 for the AF/SR classification scheme involving the combination of all features are in line with the results obtained in the literature.

With an overall accuracy of 95.0% (specificity of 92.8% and sensitivity of 96.2%), the classification performance for AF against other events (SR and VA) is encouraging (see Table 5.5). Besides, our study is among the first studies to report the performance of classification between AF and VA using PPG signals. Yet, this aspect is crucial in realistic situations as ventricular premature contractions are common in the general population. The AF/VA classification is more challenging than the AF/SR one because both arrhythmias are characterized by increased heart rate variability and disorganization. Specificity was slightly inferior (88.7%), indicating a larger number of false positive, but overall the results remained very encouraging as shown in Table 5.4.

Regarding the relevance of the proposed features, Table 5.2 shows that the best performing PPG-wave features to discriminate between SR and AF were the AOI, the VSPD and the PE, with AUC values of 0.92, 0.93 and 0.92, respectively. With an AUC of 0.60, the SPI had the lowest performance to discriminate between AF and SR and, surprisingly, larger SPI values were obtained during the AF than during SR (see Figure 5.7). As the SPI measures how well a signal can be represented by a single frequency component, the explanation for the low SPI values during SR periods may be the strong presence of the first harmonic, as shown in the PSD (see example in Figure 5.6(b)). However, the SPI was the best PPG-wave feature to discriminate between AF and VA. More generally, table 5.4 indicates that the classification of AF and VA events using only IBI-based features led to poor results, while a clear superiority of the PPG-wave features

can be noticed.

With an overall DER of 27%, the detection of heartbeats on the PPG signal was not really satisfying (see Table 5.7). As anticipated by our observations, the heartbeats were better detected during SR periods (DER of 5%) than during the arrhythmic periods (DERs of 30%). For this reason, we believe that, in some cases, the PPG-IBI features may not be sufficient because the observed increased rhythm irregularity is not only a consequence of the arrhythmia but also the result of a poorer heartbeat detection accuracy. Hence, having some features that do not rely on any heartbeat detection appears as an advantage. For comparison purposes, the classification performance using the same IBI-features from ECG are shown in Table 5.6. A drop in performance can be noticed for the PPG-IBI features, especially for the classification of AF and VA episodes. In addition, it should be mentioned that this problem is not entirely linked to the accuracy of the algorithm used for heartbeat detection, but also related to the PPG waveform itself, and therefore the amount of blood pumped towards body periphery. For example, it has been shown that the premature ventricular beats are not always followed by a pulse on the PPG signal [63]. Figure 5.8 shows an example of a very low pulsatile activity.

The amount of motion is rather limited during cardiac ablation procedures. However, motion artifacts can be of real concern when performing ambulatory monitoring with this kind of PPG device. Different approaches have been investigated so far to reconstruct clean PPG signals using adaptive cancellation of motion artifacts [42]. However, the efficiency of these methods in the presence of cardiac arrhythmias has never been demonstrated and thus, currently, the most reliable solution is the removal of the PPG segments corrupted by motion artifacts. Due to the possible resemblance between artifacts and arrhythmias, the detection of such segments can not exclusively rely on the PPG signals. Therefore, the presence of an accelerometer sensor on the wrist device is valuable.

It is important to mention that this study has some limitations. Firstly, the unbalanced dataset can be pointed out. Indeed, SR periods were rare compared to arrhythmic periods because the patients studied were suffering from various arrhythmias. Such unbalanced dataset can bias the outcome of classification. In addition, there was no patient with both AF and SR epochs. Secondly, the selected epoch duration was rather short (10 seconds). As a consequence, the number of analyzed epochs was larger but consisted of less independent epochs. However, having a short epoch duration is an advantage to identify possible short paroxysmal events. Finally, it should be mentioned that an additional feature selection step might improve the results.

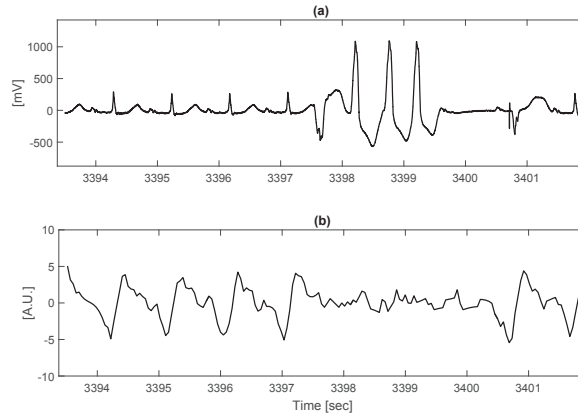


Figure 5.8: Example of a VA episode. (a) ECG (lead II); (b) PPG signal.

5.5 Conclusion

In this chapter, the detection of AF using a wrist-type PPG device was investigated. This preliminary study highlights the potential role of wrist-type PPG devices to screen AF with a reasonable accuracy. Our analysis shows that the subtle changes occurring in the PPG waveform during AF can be exploited to develop more robust classification schemes. We proposed a set of PPG-wave features that are innovative in the context of PPG processing and we showed their potential to be used as complementary measures to PPG-IBI features, especially when VA episodes are also present.

A Multimodal Approach to Reduce False Arrhythmia Alarms in the ICU

6

6.1 Introduction¹

6.1.1 Motivations

High false alarm (FA) rates are a persistent concern in the Intensive Care Unit (ICU), as FA rates up to 86% have been reported [76]. Limited performance of ICU monitoring devices results in the desensitization of the medical staff and longer response times, which can have severe repercussions. In addition, the noise disturbances that are induced may lead to patient sleep deprivation [77].

The potential of PPG to estimate HR during physical exercise and to detect AF was discussed in the previous chapters. Performance evaluation was achieved through a comparison between the results obtained using the reference sensor (i.e. the ECG) and the ones obtained using the PPG. In the present chapter, the information provided by the pulsatile signals is not used from the perspective of replacing the ECG, but as a complementary information. The physiological parameters derived from the PPG and arterial blood pressure (ABP) signals are incorporated into a multi-modal processing scheme, in order to improve the outcome compared to the case when ECG alone is used to trigger alarms. This multimodal approach aiming at reducing the number of FAs in the ICU was developed in the context of the PhysioNet/CinC challenge 2015.

This chapter is organized as follows. In Section 6.2, the methods used to process the different waveforms and determine alarms validity are presented. The results obtained are presented in Section 6.3 and discussed in Section 6.4. These results are the object of a conference paper [78] and a journal paper [79].

6.1.2 State-of-the-art

Artifacts and momentary fluctuations in the signals are the main causes for high FA rates. Studies based on heart rate (HR) trend analysis have been proposed in order to suppress FAs. An alternative approach based on a two-stage recursive median filter was proposed in [80]. The

1. The study presented in this chapter is the result of a joint collaboration with Sasan Yazdani (ASPG), which in turn will be also published in another dissertation entitled “*Novel Low Complexity Biomedical Signal Processing Techniques for Online Applications*”, thesis director Dr. J.-M. Vesin.

resulting proportion of true alarms increased from 12% to 49% during postoperative haemodynamic monitoring of cardiac patients. Several other methods have been proposed to eliminate FAs by means of multimodal signal processing. Aboukhalil *et al.* showed that the use of ABP alongside ECG can lead to an overall FA suppression of 59.7% on various alarm types, while preserving the true alarm rates, except in case of ventricular tachycardia [81].

In addition to the processing of cardiovascular signals from independent sources, the automated identification of low-quality signals and the development of signal quality indexes (SQIs) can contribute to the improvement of the decision-making process to determine the validity of an alarm [82–85]. Sun *et al.* used several features such as HR, heartbeat duration, systolic and diastolic blood pressure values to detect anomalies in the ABP waveforms and compute a signal abnormality index [82]. Another ABP quality assessment method based on fuzzy logic was proposed by Zong *et al.* [83]. These two approaches were combined by Li and Clifford to develop an ABP SQI, which was then employed to reduce the number of FAs in the ICU, resulting in FA reduction rates of 74% and 53% for extreme bradycardia and extreme tachycardia, respectively [84]. The same authors developed a more complete processing scheme by including photoplethysmogram (PPG) signals [86] and developing a PPG SQI [85]. This scheme was evaluated on a database composed of different types of life-threatening arrhythmia alarms, resulting in FA suppression rates of 86.4% for asystole, 100% for extreme bradycardia and 27.8% for extreme tachycardia, while preserving true alarms. For ventricular tachycardia alarms an FA suppression rate of 30% was achieved, with a true alarm suppression rate below 1%. The quality of ECG waveforms also plays an important role. Indeed, as arrhythmia alarms in the ICU are mainly triggered by the ECGs, low-quality ECG waveforms are associated with increases in FA rate. Behar *et al.* developed a complete scheme based on machine learning to compute signal quality indexes (SQIs) for ECG waveforms [87]. In this study, the authors pointed out the challenges linked to the evaluation of ECG waveform quality in case of abnormal rhythms and proposed to train classifiers independently for the different types of rhythms.

FA suppression in case of ventricular arrhythmia alarms can be especially arduous. Previous studies have emphasized the difficulties in classifying ventricular tachycardia and ventricular fibrillation episodes [81, 84, 88, 89]. Indeed, in these cases, accurate HR values are not sufficient to identify the presence of an arrhythmia and therefore, additional features are required to detect morphological changes in the waveforms. Moreover, there is no clear evidence yet that PPG/ABP waveforms can be used as surrogates to characterize these arrhythmias. Salas-boni *et al.* proposed a method to reduce the number of false ventricular tachycardia alarms based on ECG wavelet transform [88]. In this work, the ECG waveform was first decomposed in three sub-signals using multi-level wavelet transform. Then, different features were extracted from the sub-signals and machine learning was used to determine the validity of the alarms. An FA suppression of 21% on the PhysioNet MIMIC II dataset was achieved. Regarding the classification of ventricular fibrillation episodes, a study showed that an accuracy of 96.3% could be achieved using two features derived from ECG waveforms after the removal of noisy segments [89].

This chapter presents a new approach to lower the incidence of FAs using information from independent sources, namely ECG, ABP and PPG. As shown in Figure 6.1, four main steps are involved. Quality assessment is first performed on the available pulsatile waveforms. Then, according to the alarm type, various features are extracted from the waveforms within the window of interest. Finally the decision-making process determines whether a true arrhythmia has taken place when the alarm was triggered. Based on the quality of available signals, heart rate is either estimated from the pulsatile waveforms using an adaptive frequency tracking algorithm and/or computed from the ECG using an adaptive mathematical morphology approach. Furthermore, we propose a supplementary measure based on the spectral purity of the ECGs to determine whether a ventricular tachycardia or flutter/fibrillation arrhythmia has taken place. The proposed method is then evaluated on the PhysioNet/CinC Challenge 2015 database.

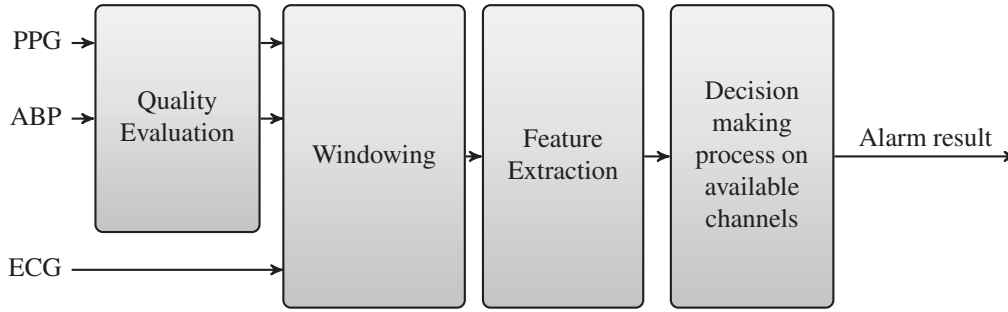


Figure 6.1: The general framework to determine the validity of an alarm.

6.2 Methods

6.2.1 Data

The PhysioNet/CinC Challenge 2015 multimodal database consists of 1250 life-threatening alarm recordings, each categorized into either bradycardia, tachycardia, asystole, ventricular tachycardia or ventricular flutter/fibrillation arrhythmia. Each record contains two ECG leads and at least one pulsatile waveform (PPG and/or ABP). The nature of each alarm was manually labeled by a team of experts according to the definitions of the five arrhythmia alarm types reported in Table 6.1. The database was divided into two subsets: a “real-time” subset, for which the data were available only before the alarm was triggered and a “retrospective” subset, in which each record contains an additional 30 seconds of data following the time of the alarm.

Table 6.1: Definition of the five alarm types [90, 91].

Alarm type	Definition
Asystole	No QRS for at least four seconds.
Extreme Bradycardia	Heart rate lower than 40 bpm for five consecutive beats.
Extreme Tachycardia	Heart rate higher than 140 bpm for 17 consecutive beats.
Ventricular Tachycardia	Five or more ventricular beats with heart rate higher than 100 bpm.
Ventricular Flutter/Fibrillation	Fibrillatory, flutter, or oscillatory waveform for at least four seconds.

For each type of alarm, two figures show examples of waveforms during the 16 seconds preceding the alarm. In these selected examples, the two pulsatile waveforms were available (ABP and PPG). Figures 6.2 and 6.3 show a true and a false asystole alarm, respectively. Figures 6.4 and 6.5 illustrate a true and a false extreme bradycardia alarm, respectively. Examples of a true and a false extreme tachycardia are reported in Figures 6.6 and 6.7, respectively. Examples of a true and a false ventricular tachycardia are reported 6.8 and 6.9, respectively. Figures 6.10 and 6.11 provide examples of a true and a false extreme ventricular/flutter fibrillation alarm, respectively.

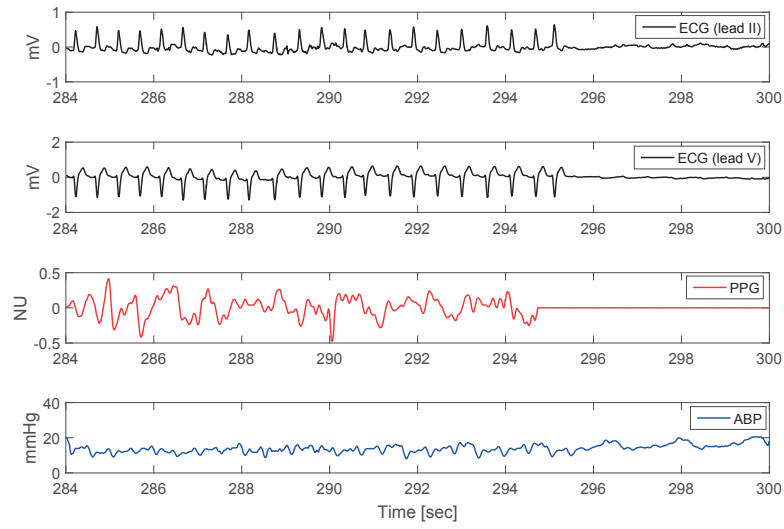


Figure 6.2: Example of a true asystole alarm (tape a142s).

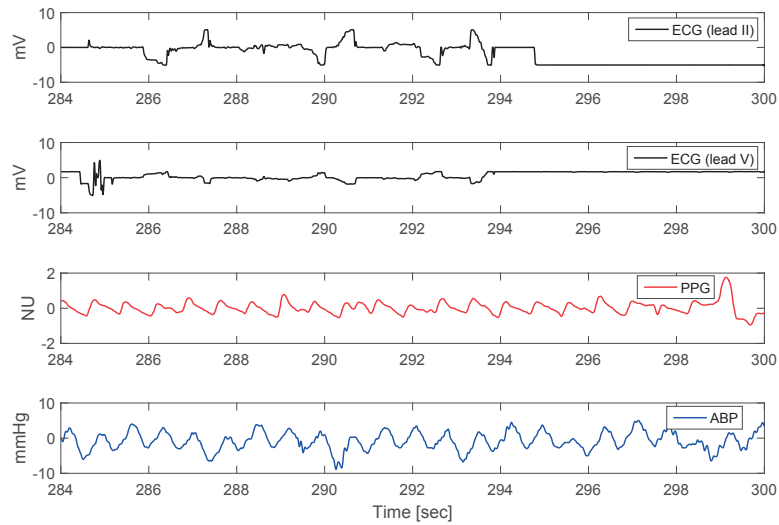


Figure 6.3: Example of a false asystole alarm (tape a134s).

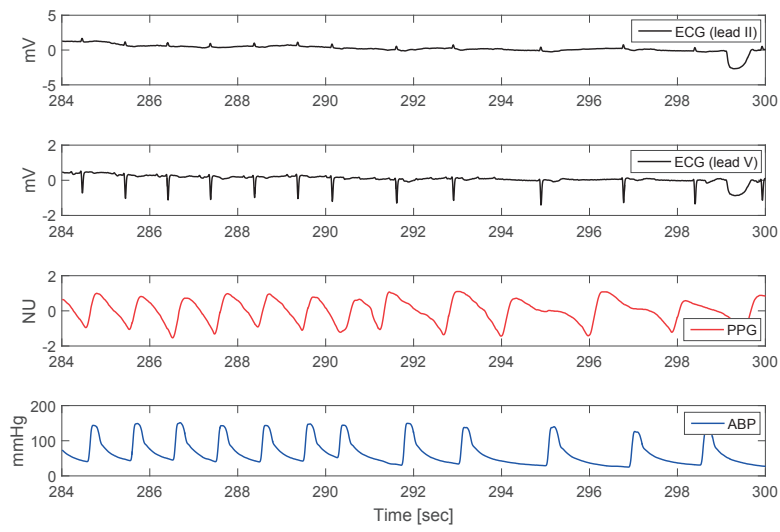


Figure 6.4: Example of a true bradycardia alarm (tape b455l).

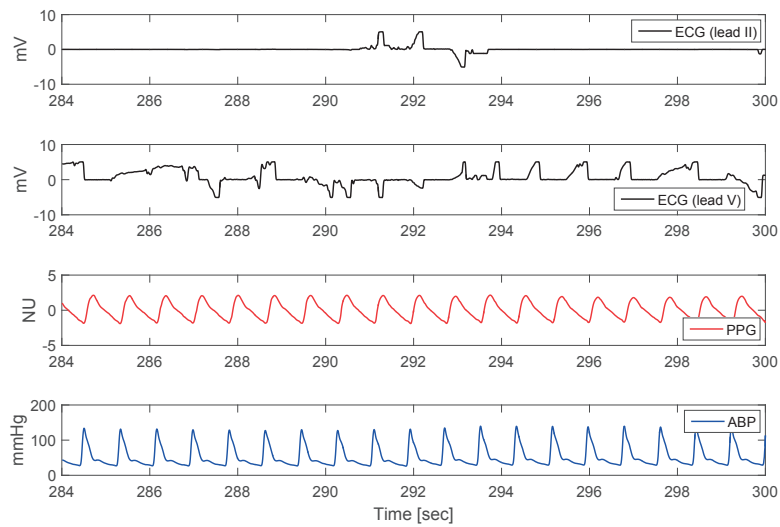


Figure 6.5: Example of a false bradycardia alarm (tape b332s).

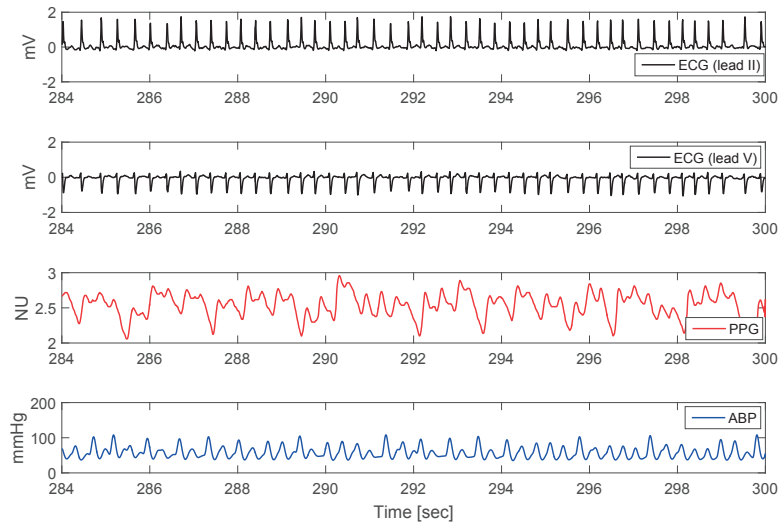


Figure 6.6: Example of a true tachycardia alarm (tape t174s).

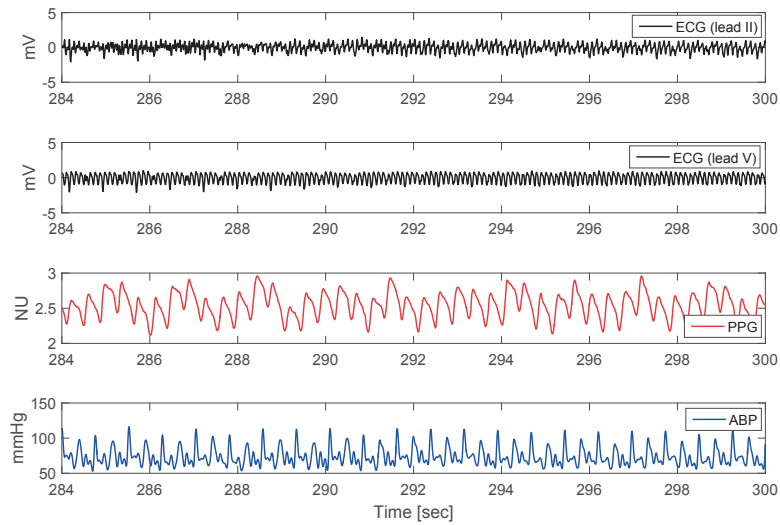


Figure 6.7: Example of a false tachycardia alarm (tape t409I).

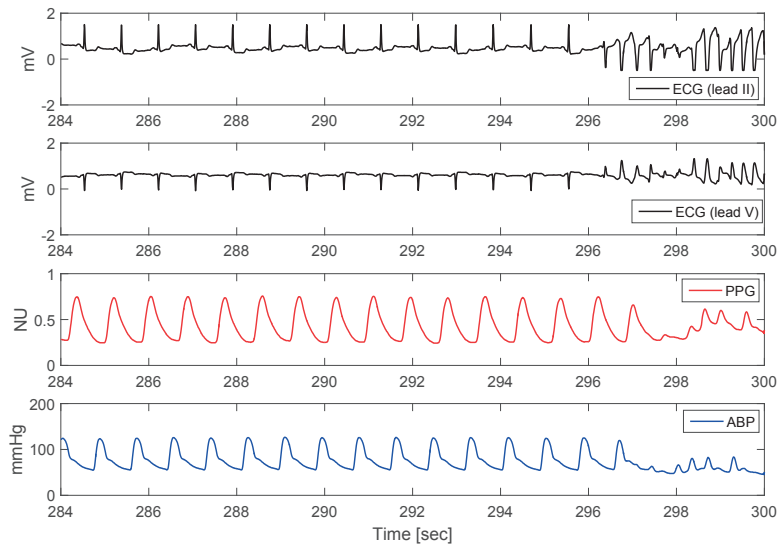


Figure 6.8: Example of a true ventricular tachycardia alarm (tape v648s).

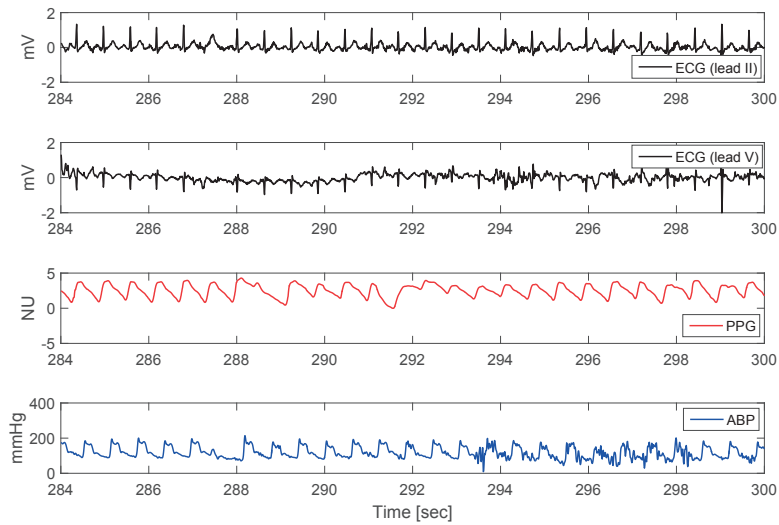


Figure 6.9: Example of a false ventricular tachycardia alarm (tape v169l).

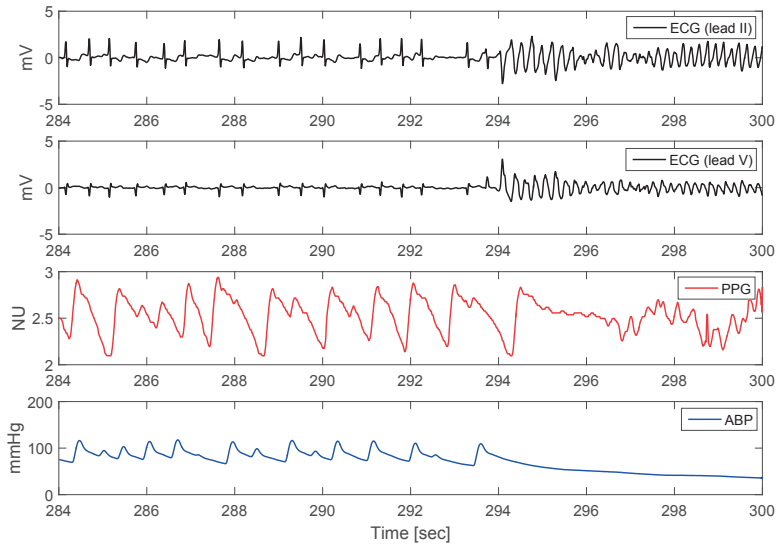


Figure 6.10: Example of a true ventricular flutter/fibrillation alarm (tape f544s).

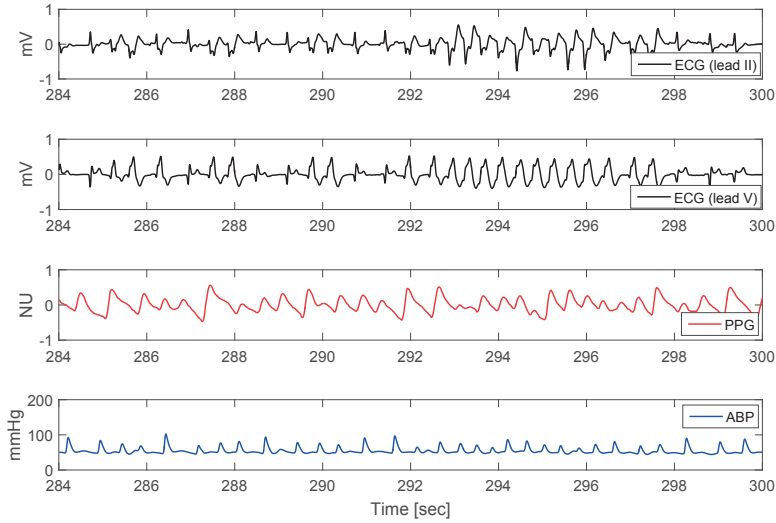


Figure 6.11: Example of a false ventricular flutter/fibrillation alarm (tape f1211).

6.2.2 ECG processing

HR estimation.

Observations of the available ECG channels in the training set indicated that they present various perturbations such as clipping of the QRS complexes, large baseline drift and high muscle

activity noise. Therefore, a robust heartbeat detection algorithm was needed in order to have a reliable FA suppression. To this end we used a QRS complex extraction algorithm proposed by Yazdani and Vesin [92, 93], based on an adaptive mathematical morphology (AMM) approach. In this setting, operators nonlinearly transform the signal of interest using a structuring element (SEI), designed to extract useful information regarding shape and size. The resulting signal is called the feature signal. Figure 6.12 illustrates the general framework of AMM.

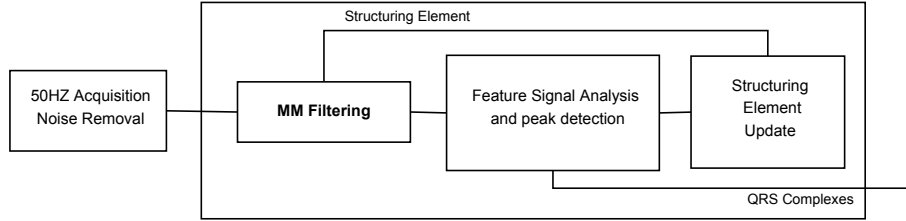


Figure 6.12: The block diagram of AMM.

The outcome of a mathematical morphology operator depends on the shape of the SEI. AMM uses a synthesized SEI, in order to extract heartbeats from the ECG. synthesized SEI comprises five fiducial points namely the onset, Q-wave, R-wave, S-wave and the offset. With a detection of a heartbeat, a series of morphological features such as the shape and size (i.e. the amplitude and the relative distance to the onset) are extracted from the beat. Then using a learning coefficient the amplitude and the location of each fiducial point in synthesized SEI are updated. The new synthesized SEI is used to extract future QRS-complexes in the ECG. This adaptation of the SEI makes AMM flexible, and makes possible a more robust QRS complex extraction against the aforementioned perturbations. Furthermore, the learning coefficient in AMM is tuned based on the change in power of two successive heartbeats. This tuning on the learning coefficient is set to avoid excessive changes on the synthesized SEI, especially when ECG presents impulsive noise or when ectopic beats such as PVC (or bigeminy and trigeminy) take place in the ECG. Figure 6.13, illustrates the performance of AMM on a tape from the training set with pathological ventricular rhythm (bigeminy). Finally, it is worth mentioning that the pacing spikes are not detected in the scheme due to their limited time duration. From a practical view point, AMM avoids excessive use of arbitrary thresholds, and entails a low computational cost, linear with respect to the length of the signal. Figures 6.14 illustrates the performance of AMM on two low-quality signals from the training dataset.

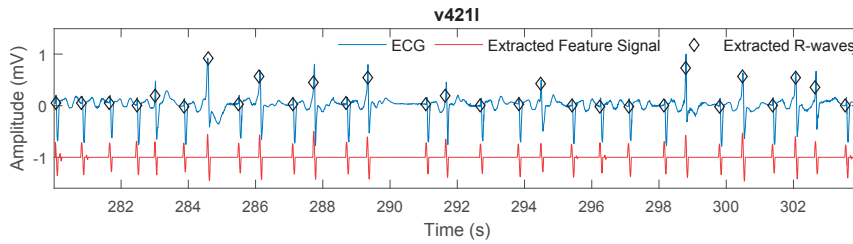


Figure 6.13: AMM performance on a tape with bigeminy (v4211).

Spectral purity index.

Defining the characteristics of ventricular arrhythmia is challenging from a signal processing point of view. More specifically, in case of ventricular tachycardia, the definition includes a con-

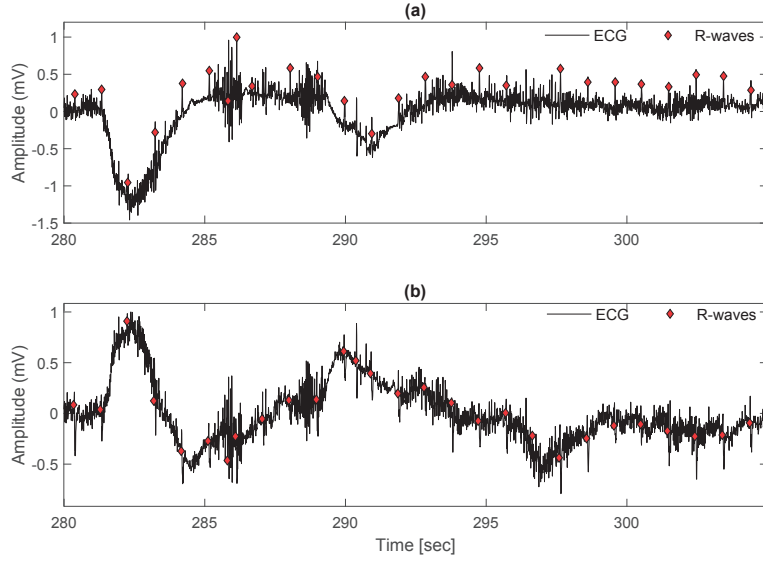


Figure 6.14: AMM performance on an ECG (tape a527l) from the training set.

straint on the rhythm (HR higher than 100 bpm) as well as morphological changes in QRS complexes (five or more ventricular beats). Although many methods have been proposed to estimate the HR, methods to identify a succession of ventricular beats are scarce. Our observations have shown that ECGs become closer to a sinusoid during ventricular tachycardia/flutter/fibrillation episodes, due to a widening of the QRS complexes. Such kind of behavior can be quantified with a measure called the spectral purity index (SPI). This measure, which ranges between zero and one, is described in Section 2.4. In this study, the SPI of the available ECG channels was measured for ventricular tachycardia and ventricular flutter/fibrillation alarms. Higher SPIs were expected for true arrhythmias. In order for the SPI to be more representative of the general changes in the signals, ECG waveforms were first smoothed by down-sampling to 35 Hz and by applying a 5-sample moving average filter. A 2-second sliding window ($L=70$) was used for SPI calculation. Figure 6.15 illustrates an example of the SPI during a true ventricular tachycardia episode.

6.2.3 PPG and ABP processing

Quality assessment.

In order to assess the quality of the PPG and ABP signals, we used respectively the *ppgSQI* and the *jSQI* algorithms, provided for the PhysioNet/CinC 2015 challenge [90, 91]. Based on the detected heartbeats, these algorithms compute the features needed to estimate signal quality. Heartbeats were detected using the algorithm described by Arberet *et al.* [94]. The resulting signal quality indexes (SQIs), that ranged between zero and one, determined whether PPG/ABP waveforms should be analyzed.

HR estimation.

In order to suppress FAs and preserve true alarms, it is necessary to have an accurate HR extraction. However, a precise HR estimation has proven to be difficult when signal quality is low.

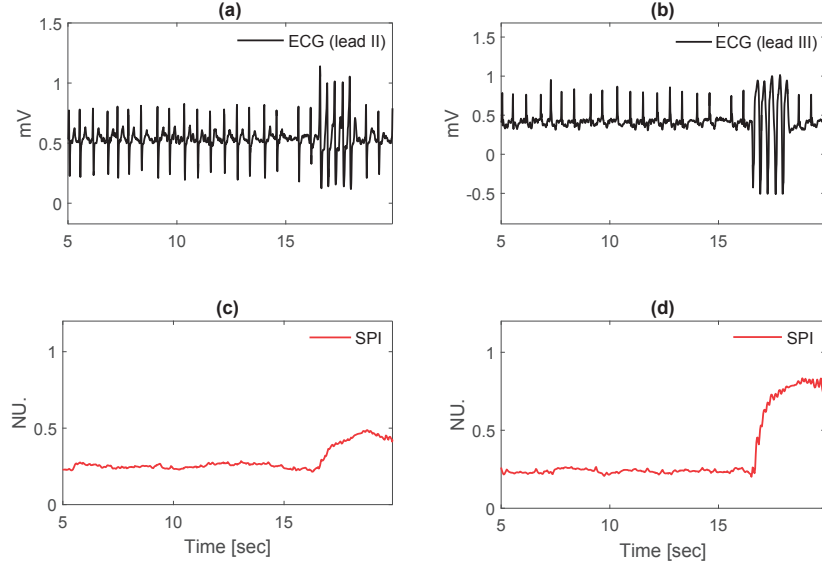


Figure 6.15: Example of SPI for a true ventricular tachycardia (record v194s). (a) and (b) show the two ECG waveforms; (c) and (d) represent the corresponding SPIs. In both cases, it can be observed that the SPI significantly increases when the arrhythmia takes place.

Moreover, abnormal heart rhythms are expected to be present in the waveforms. In this study, we decided to exploit the pulsatile nature of PPG and ABP waveforms by using adaptive frequency tracking algorithms, which consist of adaptive band-pass filters, to estimate HR. More specifically, the OSC-MSE-W and the SFT-W algorithms described in Sections 2.3.1 and 2.3.2 were used. An 8th-order Butterworth low-pass filter with a cutoff frequency of 5 Hz was first applied to the PPG signals. Then, the baseline of PPG and ABP signals was removed by subtracting the mean of the upper and lower signal envelopes estimated using maximum/minimum detection on a sliding window. Finally, when the SQIs of the PPG/ABP signal reached a certain threshold, the OSC-MSE-W/SFT-W algorithms were used to compute the instantaneous HR of the available signals. In order to increase the robustness of HR estimation, smoothed versions of the input signals, obtained with a moving average of length l_{win} , were also fed to the algorithm. It is especially important to attenuate the dicrotic notch and to ensure the tracking of the main oscillation and not of the first harmonic. The example illustrated in Figure 6.16 shows why it is necessary to use also the smoothed signal as an input to the adaptive frequency tracking algorithms for a precise HR estimation. In this example, the ABP signal has a lower frequency component, which is the frequency corresponding to instantaneous HR, and a higher frequency component, due to the presence of the dicrotic notch. It can be seen that, when the smoothed ABP signal is also provided to the OSC-MSE-W algorithm, the common frequency between the inputs is the frequency of interest and the HR can be accurately estimated, which is not the case when only the ABP signal is used. In order to further optimize HR estimation, the parameters required for adaptive frequency tracking, were empirically selected for each type of arrhythmia. These parameters are summarized in Table 6.2, where f_{re} indicates the re-sampling frequency of the waveforms, β is related to the bandwidth of the adaptive band-pass filter and δ is a forgetting factor. It should be noted that, for asystole alarms, a smaller δ was used in order to detect decreases in HR faster. We noticed that the complex version of the algorithm (SFT-W) was more suitable

for arrhythmia associated with low HR, i.e. asystole and bradycardia, while the OSC-MSE-W was more suitable for arrhythmia associated with high HR. Figure 6.17 shows an example of HR estimation using the OSC-MSE-W algorithm, during a true extreme tachycardia episode. As illustrated, despite the moderate quality of the waveforms, the oscillation of interest is correctly isolated by the adaptive band-pass filter and therefore, instantaneous HR can be estimated.

Table 6.2: Algorithms and selected parameters for HR estimation using PPG and/or ABP waveforms.

Arrhythmia	Algorithm	f_{re} [Hz]	l_{win} PPG [samples]	l_{win} ABP [samples]	β	δ
Asystole	SFT-W	15	3	3	0.8	0.8
Extreme Bradycardia	SFT-W	15	3,5,11	7	0.87	0.87
Extreme Tachycardia	OSC-MSE-W	35	5	7	0.89	0.9
Ventricular Tachycardia	OSC-MSE-W	35	5	7	0.89	0.9

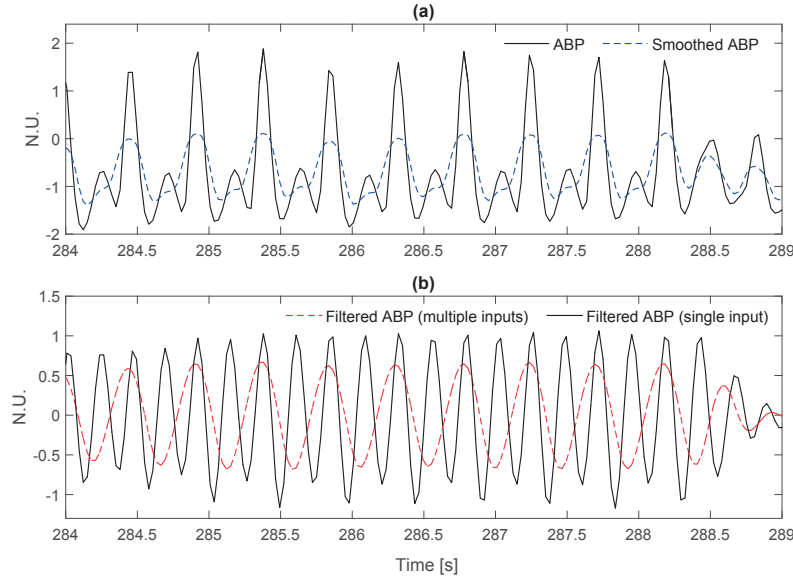


Figure 6.16: An example to illustrate the interest of using multiple inputs for adaptive frequency tracking. (a) Original (baseline removed) and smoothed ABP signals; (b) ABP signal filtered with the adaptive band-pass filter (single input to the OSC-MSE-W) and ABP signal filtered with the adaptive band-pass filter (multiple inputs to the OSC-MSE-W). The HR component is correctly extracted when ABP and smoothed ABP are used as inputs.

6.2.4 Arrhythmia alarm processing

Definition of the five alarm types (see Table 6.1) enabled us to develop processing schemes specific to each arrhythmia, which are described throughout the upcoming subsections. In each scheme, thresholds were empirically selected in order to maximize the corresponding true negative rate (TNR) while keeping true positive rate (TPR) as close as possible to 100% in the training

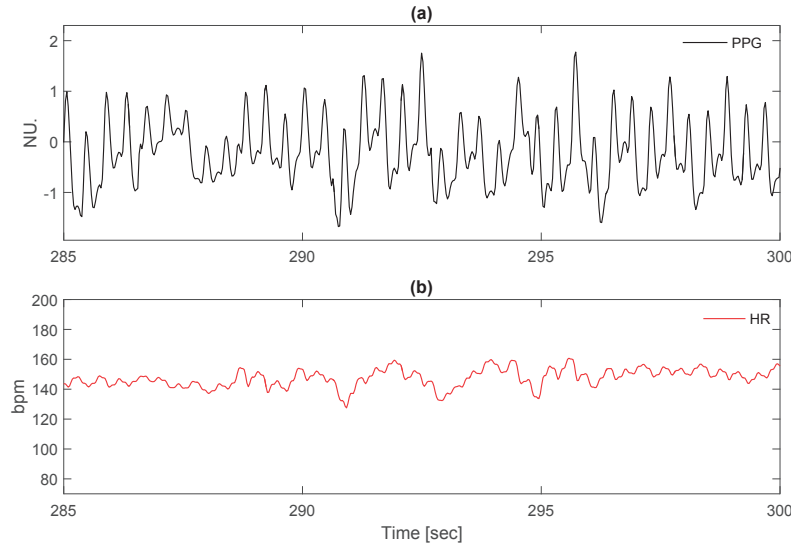


Figure 6.17: An example of HR estimation using the OSC-MSE-W algorithm during an extreme tachycardia episode. (a) PPG waveform; (b) Instantaneous HR.

dataset. It should be mentioned that, for each type of arrhythmia, the default alarm result was set to one. Instead of detecting a given arrhythmia, our algorithms rely on finding sufficient conditions to suppress the alarm.

Asystole processing.

Figure 6.18 illustrates the processing scheme for asystole. For this arrhythmia, available ECGs and high-quality pulsatile waveforms were processed separately and alarm veracity was determined by means of voting. More specifically, the alarm was suppressed if at least one processed waveform suppressed the alarm. For the retrospective setting, a few seconds of data after the alarm were also taken into account for the calculation of the HR features, i.e. the mean HR and the decrease of HR for the pulsatile waveforms and the maximum RR-interval for ECG waveforms. A linear discriminant analysis (LDA) classifier, trained on the training dataset (see Section 6.2.5), was used to compute the alarm result derived from the pulsatile waveforms. In order to have consistency with the performance metric that was used in this study (see Section 6.2.5), the false negative cost was set as five times larger than that of false positive in the LDA. The classifier was trained on the training dataset. Regarding the processing of the ECG signals, a threshold of 3.5 seconds, was put on the maximum RR-interval during the last 20 seconds before alarm was triggered (+ 5 seconds after in case retrospective data were available) to compute the ECG alarm results.

Extreme bradycardia processing.

Two different processing schemes are possible in case of extreme bradycardia alarm, depending on the quality of the waveforms (Figure 6.19). When at least one of the pulsatile waveforms had an acceptable quality (SQI higher than 0.5), alarm validity was determined based on a threshold of 54 bpm on the mean and median HR values extracted from the pulsatile waveforms during the last five seconds before the alarm was triggered. On the other hand, when the SQIs of PPG

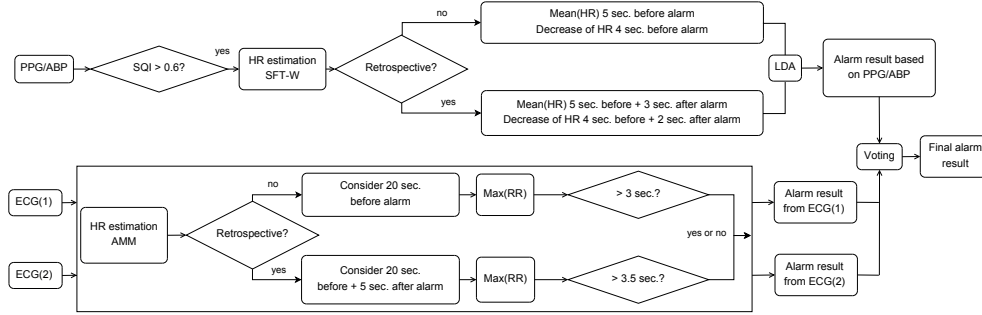


Figure 6.18: Asystole alarm processing.

and ABP waveforms were both below 0.5, the ECG waveforms were analyzed. A threshold of 40 bpm was used on the minimum HR derived from five consecutive beats to discriminate between true and false alarms. A period of 16 seconds before the alarm was considered and, if available, 5 seconds after the alarm were also taken into account.

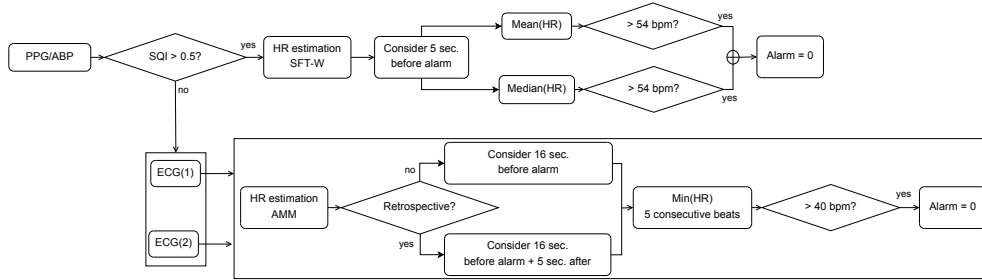


Figure 6.19: Extreme bradycardia alarm processing.

Extreme tachycardia processing.

In case of extreme tachycardia alarm, only the pulsatile waveforms were processed, as shown in Figure 6.20. The decision about alarm veracity was made from the mean and maximum HR values during the last four seconds preceding the alarm. More precisely, the alarm was suppressed if the maximum HR was below 115 bpm while the mean HR was below 100 bpm. For this arrhythmia, ECG waveforms were not analyzed.

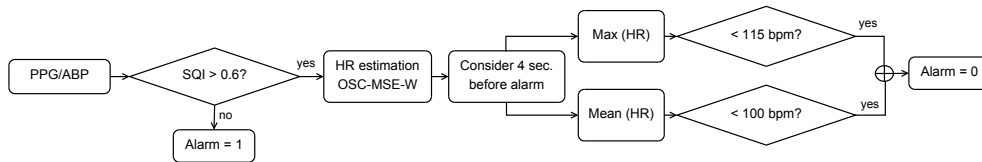


Figure 6.20: Extreme tachycardia alarm processing.

Ventricular flutter/fibrillation processing.

As shown in Figure 6.21, a single feature, the maximum average SPI on a 3-second sliding window (2-second overlap), was used to suppress false ventricular flutter/fibrillation alarms.

More precisely, the alarm was suppressed if this SPI feature was smaller than 0.63 during the 15 seconds preceding the alarm. Pulsatile waveforms were not processed for this arrhythmia.

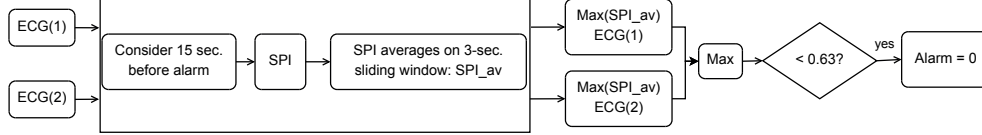


Figure 6.21: Ventricular flutter/fibrillation alarm processing.

Ventricular tachycardia processing.

The processing of ventricular tachycardia alarms was the most challenging one (Figure 6.22). On one hand, pulsatile waveforms were used to compute HR and suppress the alarm if the HR ranged between 60 and 90 bpm. On the other hand, the ECG SPI was employed to detected morphological changes during ventricular contractions. It should be noted that pulsatile waveforms and ECGs were processed independently, and that both could lead to the suppression of the alarm. The first feature was the maximum of the average SPI on a 3-second sliding window (2-second overlap) during the 20 seconds preceding the alarm. The maximum difference between every other value of these SPI averages were used to compute a second feature reflecting SPI increase. Three different conditions led to alarm suppression based on the SPI features described above. Either the maximum SPI increase had to be below 0.012, or the maximum SPI average was below 0.25, or both the maximum SPI average and the maximum SPI increase were below 0.36 and 0.2, respectively.

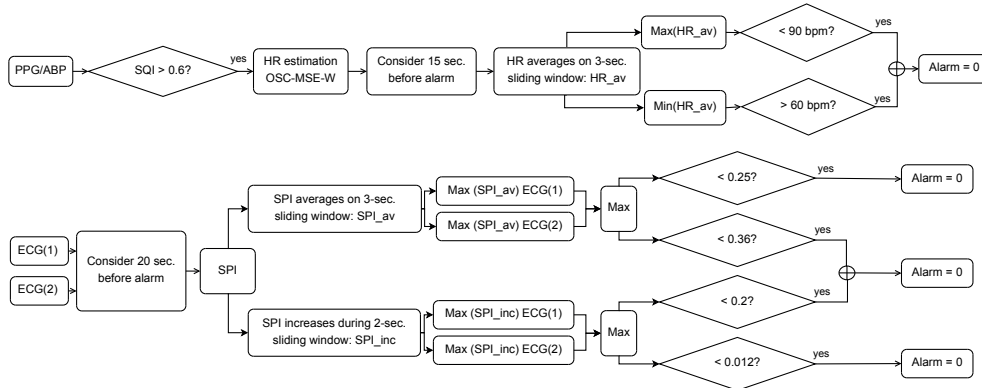


Figure 6.22: Ventricular tachycardia alarm processing.

6.2.5 Evaluation metrics

A training set of 750 records was provided for this challenge. Then, the performance of the developed scheme was evaluated on a hidden test set of 500 records (extended to 750 records for the second phase). It is important to eliminate false alarms while preserving true alarms, since the life of the patient is threatened if true alarms are suppressed. To enforce this, a specific score function was selected for this challenge [90, 91]. This score is a function of the number of true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN), computed by

the following equation:

$$\text{Score} = \frac{100 \cdot (\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + 5 \cdot \text{FN})} \quad (6.1)$$

As shown in Eq. 6.1, FN alarms are penalized five times more than FP alarms. In addition to this score, TPRs and TNRs were reported for the different alarm types (see Appendix A for an introduction about evaluation of classifiers).

6.3 Results

After feature extraction and training on the available training set, our scheme was first evaluated on a test set of 500 records, and was ranked as the second best method with an overall score of 73. For the second phase of this challenge, we focused on optimizing the parameters that were used by the different algorithms of our scheme. At the same time, different voting strategies were studied in order to find the best way to aggregate the results provided by each processing component. It is worth mentioning that, for this phase, an additional dataset of 250 records was added to the evaluation dataset. For the real-time test dataset, the proposed framework achieved a TPR of 94%, a TNR of 77%, and an overall score of 76.11. Regarding the retrospective subset, a TPR of 99% and a TNR of 80% were achieved, with an overall score of 85.04 [95]. This scheme was ranked fifth in the real-time event and first in the retrospective event [90, 91]. A follow-up phase was also considered for this challenge. After applying minor changes to the decision-making process of the asystole and extreme tachycardia alarms, an improvement of one percent in TPR was achieved for the real time event, resulting in a real-time score of 77.07. More details on the performance of the proposed scheme on the test and the training datasets can be found in Tables 6.3 and 6.4, respectively.

Table 6.3: Results obtained on the hidden-test dataset.

Arrhythmia	Phase I			Phase II			Follow-up		
	TPR [%]	TNR [%]	Score	TPR [%]	TNR [%]	Score	TPR [%]	TNR [%]	Score
Asystole	92	78	76.42	83	88	81.44	94	85	84.28
Extreme brady.	96	66	73.53	100	71	82.47	100	71	82.47
Extreme tachy.	96	60	80.00	97	60	86.18	97	80	86.99
Ventricular flut./fib.	83	88	79.55	89	94	87.10	89	94	87.10
Ventricular tachy.	93	65	67.38	94	72	72.75	94	72	72.75
Real-time	93	65	68.15	94	77	76.11	95	76	77.07
Retrospective	95	77	77.82	99	80	85.04	99	80	85.04

Table 6.4: Results for the training set.

Arrhythmia	TP [%]	FP [%]	FN [%]	TN [%]	TPR [%]	TNR [%]	Score
Asystole	18	12	0	70	100	85	87.70
Extreme bradycardia	51	18	1	30	98	63	77.49
Extreme tachycardia	94	3	0	4	100	55	97.10
Ventricular flutter/fibrillation	10	16	0	74	100	83	84.48
Ventricular tachycardia	22	19	4	55	85	75	67.45
Average	39	13	1	47	98	78	82.31
Gross	37	14	2	47	95	76	77.88

6.4 Discussion

In this challenge, the highest score in both real-time and retrospective events was achieved by a meta-algorithm, which uses a voting system on the 13 best performing entries (see Appendix A for a short introduction about classifier fusion). Table 6.5 compares the results achieved by our scheme with the meta-algorithm developed by the challenge organizers. More details about the top scoring entries from the other participants can be found in [90, 91]. Regardless of our final scores, our main goal was to achieve the highest TPR since the applicability of the developed method in the ICU would be most acceptable with a TPR of 100%. This was reflected in the final results, as our approach reached the highest TPR for both real-time and retrospective events [90].

Table 6.5: Comparison between our scheme and the voting algorithm.

	Real-time subset			Retrospective subset		
	TPR	TNR	Score	TPR	TNR	Score
Voting algorithm [90]	94%	90%	84.26	94%	94%	87.03
Our scheme	95%	76%	77.07	99%	80%	85.04

We believe that quality assessment of the signals plays an important role in FA suppression in the ICU. Indeed, in our case, the use of the PPG/ABP SQI was necessary to obtain satisfying results. Table 6.6 shows the percentage of records that were analyzed based on the SQI thresholds set in our processing scheme. It should be noted that for extreme tachycardia alarms, only 51% of the records had quality acceptable enough to be analyzed. Therefore, excluding 49% of signals undoubtedly affects the final TPR and TNR that can be achieved for this arrhythmia, as shown by the low TNR on the training set reported in Table 6.4. The use of an ECG quality assessment method was also investigated in this study. As mentioned before, the ECGs in the database usually present various perturbations. Therefore, during the training phase of this study, we developed an ECG quality assessment classifier, in order to categorize the ECGs into low, moderate, and high quality classes. To this end, all ECGs in the dataset were manually labeled by an operator and various features were extracted from them. Initially, third and fourth moment based features, i.e. respectively skewness and kurtosis, and ECG QRS- and baseline-power features were extracted from the tapes, as proposed in [87]. Then ECGs were processed by AMM [92] and several QRS complex-driven features were extracted from the output of AMM. More specifically, features such as maximum, minimum, average duration of QRS complexes and RR-intervals were extracted from the signals. Subsequently, a support vector machine based ECG signal quality assessment classifier was trained and validated (10-fold cross-validation). The classification accuracy of the trained classifier was 81.6% against the training set. However, when this ECG quality assessment was incorporated into AMM, i.e. only ECGs with high enough quality would be processed, a performance decrease was observed in the training set. For this reason, the use of ECG quality assessment was discontinued for this challenge.

Table 6.6: Percentage of processed waveforms for each type of alarm in the training set. The last column indicates the percentage of records for which at least one waveform was processed.

% of processed records	ABP	PPG	PPG or ABP	ECG	At least one waveform
Asystole	35%	45%	80%	100%	100%
Extreme bradycardia	40%	64%	80%	20%	100%
Extreme tachycardia	29%	26%	51%	0%	51%
Ventricular flutter/fib.	0%	0%	0%	100%	100%
Ventricular tachycardia	45%	52%	78%	100%	100%

Based on the results presented above, it seems that adaptive frequency tracking is an appropriate technique to obtain instantaneous HR from the pulsatile waveforms. This method does not provide any information about heart rate variability, but results suggest that precise HR averages and trends are most of the time sufficient to determine the validity of the alarm. Moreover, adaptive frequency tracking enabled us to process moderate-quality waveforms, and as a result increased the numbers of potential FA candidates for suppression. Furthermore, information from ABP and PPG can be easily combined using this technique.

We found out in this study that the elimination of false ventricular tachycardia alarms only on the basis of PPG/ABP waveforms is very difficult. Our observations showed that ventricular tachycardia episodes sometimes induce amplitude decreases in the pulsatile signals. However, as illustrated in Figure 6.23, this effect was not consistent over records. In this figure, the PPG waveforms, as well as the ECG waveforms are displayed for two records corresponding to true ventricular tachycardia episodes. It can be noted that the pulse amplitudes decrease during the ventricular beats for the first record (highlighted by the ellipse), but the same behavior can not be observed for the second record. For this reason, PPG/ABP amplitude changes were excluded from our scheme. This phenomenon had been reported by [63]. In this work the pulse amplitude, among other features, was used to detect ventricular premature beats. The authors noted that different pulse patterns are possible in the PPG signal when a premature ventricular beat occurs, depending on blood pumping efficiency. Finally, as a compromise, we decided to rely on pulsatile waveforms for HR estimation and on ECGs for the detection of the morphological changes. To the best of our knowledge, it is the first time SPI is used to characterize the morphological changes related to ventricular arrhythmia. This measure undoubtedly has a potential to be used for ventricular tachycardia and ventricular flutter/fibrillation arrhythmias, as confirmed by the good results achieved with a very limited number of features. Moreover, the computation of this index can be done on-line in the time domain, at a very low cost.

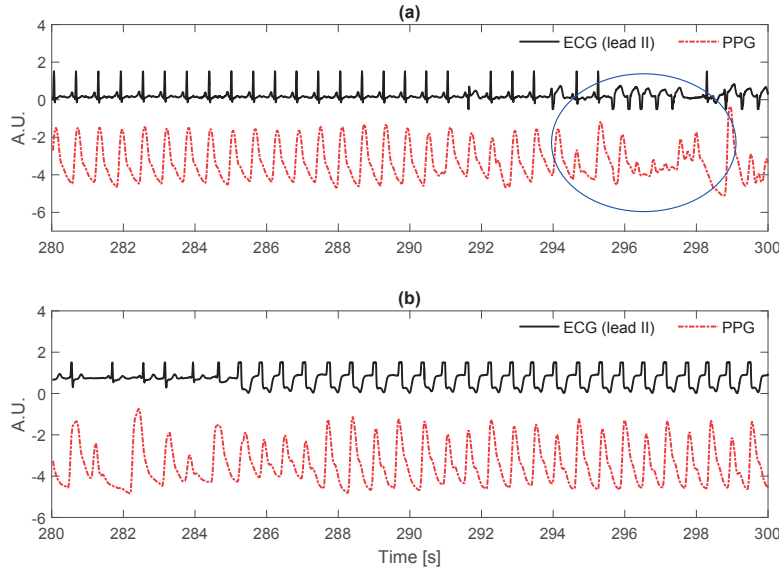


Figure 6.23: Two examples of PPG signal behavior during true ventricular tachycardia episodes. (a) Record v628s. A decrease in pulse amplitude can be observed in the PPG signal during the arrhythmia; (b) Record v159l. No decrease in PPG pulse amplitude.

The decision-making process of our scheme is very straightforward, as it mainly uses simple thresholds on a few features to eliminate FAs. In order to find the optimal way to aggregate the results provided by different components of our scheme, both learner fusion and learner selection methods were studied. As illustrated by the arrhythmia alarm processing schemes (Figures 6.18 - 6.22), for asystole and ventricular tachycardia alarms, a fusion-based method leads to the best results. For extreme bradycardia, the alarm verification procedure is chosen based on the threshold of 0.5 on the PPG/ABP SQI, which represents the learner selection that was incorporated for this arrhythmia. For extreme tachycardia and ventricular flutter/fibrillation, results showed that the use of pulsatile waveforms alone and ECG waveforms alone, respectively, were more effective. It should also be noted that, compared to the exact alarm definitions (see Table 6.1), most HR thresholds in our method were set to have a certain degree of tolerance, which was necessary to avoid the suppression of true alarm. Using a decision-making process based on sophisticated machine learning algorithms is debatable in our case. Indeed, machine learning-based approaches have already proven their efficiency, as shown by the results from the other participants in the challenge [90]. However, our limited number of features allowed us to keep our scheme simpler while rendering a physiological interpretation possible. The similarities between the performance achieved on training and test datasets confirm the relevance of the empirically chosen parameters (see Tables 6.3 and 6.4), and suggest that overfitting was avoided despite the limited number of available records for some types of arrhythmia. It is noteworthy to mention that, during the development of our method for the processing of ventricular tachycardia, we tried to replace the thresholds on SPI derived features (see Figure 6.22) by an LDA classifier. However, despite the better results obtained on the training set, lower performance were achieved in the test dataset, as shown in Table 6.7.

Table 6.7: Linear discriminant analysis for ventricular tachycardia alarms.

	Training dataset			Test dataset		
	TPR [%]	TNR [%]	Score	TPR [%]	TNR [%]	Score
With LDA	90	77	72.83	81	74	64.83
No LDA	85	75	67.45	94	72	72.75

6.5 Conclusion

The study presented in this chapter confirms that the use of cardiovascular signals from independent sources is of great interest to reduce the number of FAs in the ICU. We propose an innovative approach, based on the capture of specific signal behavior that occur during arrhythmia episodes, in addition to the traditional HR measures derived from the ECGs and pulsatile waveforms. The efficiency of our approach was demonstrated on the PhysioNet/CinC Challenge 2015 dataset, for which TNRs of 76/80% were achieved, with corresponding TPRs of 95/99% for the real-time and retrospective subsets, respectively.

Part III

Processing of Video-Derived Photoplethysmographic Signals

Introduction to Imaging Photoplethysmography

7

7.1 Principle

Imaging photoplethysmography (iPPG) was introduced for the first time in 2008 [5]. This technique enables the remote measurement of cardiac pulses. Similarly to PPG, iPPG is an optical-based method. However, iPPG differs from the usual PPG technique in various aspects. While PPG requires an illumination source at a dedicated wavelength, iPPG is a non-contact technique that uses ambient light as only illumination source. The camera sensor, which can be a low-cost commercial digital camera, is typically placed 1-2 meters away from the subject. The iPPG signals are derived from temporal skin color variations of a skin region. This region of interest (ROI) is generally selected on the subject's face. According to the skin reflection model described in [96], the signal measured by the camera sensor is a mixture of specular and diffuse reflection. As shown in Figure 7.1, the pulsatile information is contained in the diffuse reflection. Unlike PPG, which measures the blood volume changes from a single site, iPPG measures these variations by averaging the changes in a relatively large ROI.

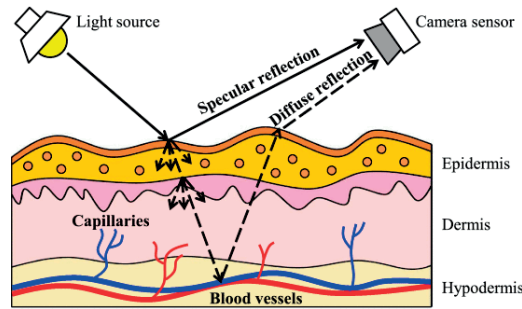


Figure 7.1: Skin reflection model. Image from [96], © 2016 IEEE.

Different kinds of cameras have been previously used for iPPG applications. In [97], iPPG signals from a webcam and a high-speed CMOS color camera were compared. In this study, the authors observed similar capabilities in measuring HR from both systems. Besides the selected camera, the video compression is a very important parameter. McDuff *et al.* compared a set of commonly used video compression algorithms (x264 and x265) at different constant rate factors (“level of compression”) [98]. The analysis was performed on two data subsets, namely a stationary task and a random motion task. The authors concluded that compression, even at

low constant rate factors, considerably degrades the blood volume pulse signal-to noise ratio. Moreover, this effect is worsened in the presence of additional noise sources such as large head movements. It is therefore important to analyze the raw video data.

Various techniques, reviewed in the next section, have been previously explored in the literature to derive iPPG signals from video sequences. It should be mentioned that differing terminologies are used in the literature to refer to the iPPG technology, including among others remote photoplethysmography (rPPG) and video photoplethysmography (vPPG). The term iPPG is used throughout this dissertation.

7.2 Review of the iPPG-signal derivation methods

Spatial averaging

In [5], Verkruysse *et al.* showed for the first time that PPG signals can be remotely measured from a subject's face with a simple camera using normal ambient light as illumination source. In their study, a digital camera placed 1-2 meters away from the subjects was used to record video-sequences of their face. The illumination conditions consisted of a combination of daylight and normal artificial fluorescent light. The ROI was selected on the subject's face and raw signals were computed as the average of all pixel intensity values within the ROI at each frame and for each channel of the RGB image (namely RED, GREEN and BLUE channels). The raw signals were then band-pass filtered between 0.8 and 6 Hz (48-360 bpm), in order to detrend them and remove very high frequency components. Time-frequency and frequency domain analyses were performed on the raw and band-pass filtered iPPG signals. The main conclusions of this preliminary study can be summarized as follows: 1) the HR estimate was found to be in excellent agreement with the one provided by the reference sensor, 2) averaging pixels within a ROI leads to higher SNR, 3) the green channel is the best channel, and 4) the respiratory rate can also be determined. The authors also emphasized the limitations related to motion artifacts, as well as to the noise in the pixel values generated by the charge coupled device (CCD) camera. Figure 7.2 illustrates the basic methodology to recover iPPG signal from the video sequences.

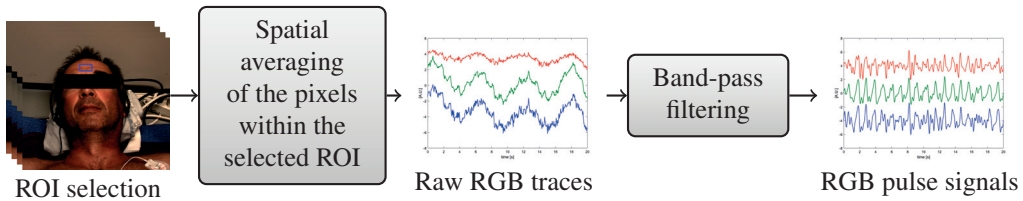


Figure 7.2: Basic principle of iPPG signal derivation, as introduced in [5].

Blind-source separation based methods

Based on the observations that 1) the iPPG technique is very sensitive to noise and 2) the frequency band of the noise overlaps with the frequency band of the physiological signals of interest, researchers decided to investigate blind-source separation (BSS) methods in order to retrieve the blood volume pulse [99–102]. BSS methods allow the separation of the sources from a set of observed mixed signals, for which the mixing process is unknown. The general methodology for iPPG-signal derivation using BSS is illustrated in Figure 7.3. In the BSS-based

processing scheme proposed in [99], the raw RGB traces were first normalized as follows:

$$x_{norm}(t) = \frac{x(t) - \mu}{\sigma} \quad (7.1)$$

where μ is the mean and σ the standard deviation. Then, independent component analysis (ICA) was used to decompose the normalized RGB traces into three independent source signals. A time-window of 30 seconds was used for this purpose. The information contained in the signals before and after ICA was compared, and the authors showed that the HR estimated using ICA was more accurate compared to the HR estimated from the raw traces, especially for sequences with motion artifacts. In [100], the same authors extended their ICA-based methodology to quantify additional physiological parameters, namely respiratory rate and HRV parameters (LF, HF and LF/HF)¹. For this purpose, a supplementary processing step was added: the selected source signal was smoothed and band-pass filtered between 0.7 and 7 Hz. In order to enhance the precision of heartbeat detection, the signal was interpolated at 256 Hz. For immobile subjects, a correlation of 0.94 was reported between the reference interbeat intervals (IBI) and the iPPG-IBI.

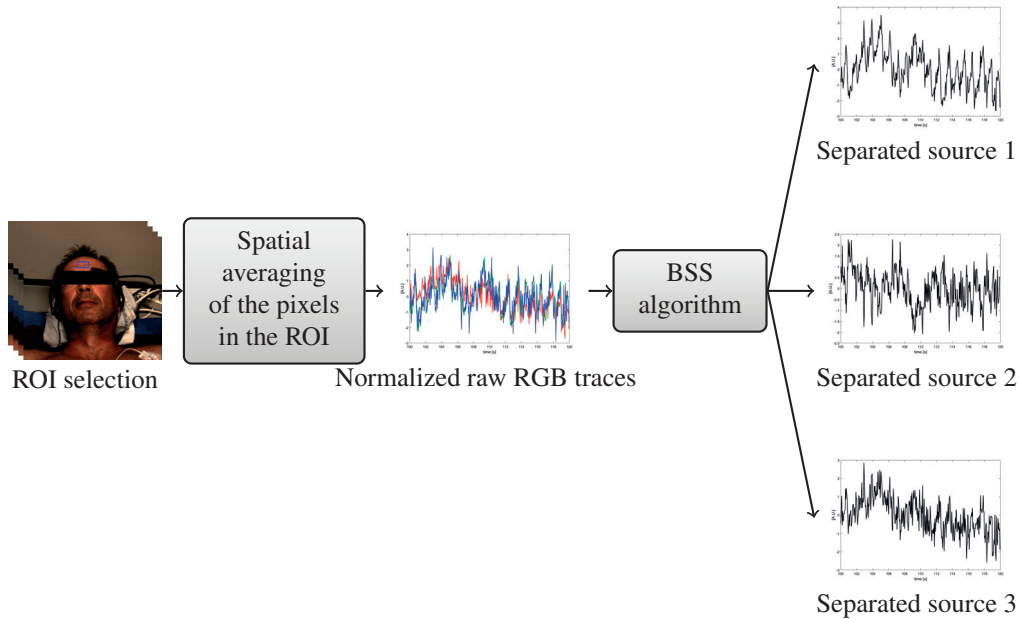


Figure 7.3: Principle of BSS-based iPPG signal derivation methods.

Model-based methods

Given the drawback induced by the component selection step in the BBS methods and the fact that these methods are not optimal for real-time applications, other iPPG signal derivation methods have been proposed in the literature. A chrominance-based iPPG method was introduced in [103], in which the iPPG signal is reconstructed as a linear combination of the RGB color channels. The authors suggested that the specular reflection component, which is not affected by blood volume changes, as well as the non-local intensity variations, were the main causes of distortion. In their model, a constant standardized skin-color is assumed to normalize the signal

1. LF: power in the low frequencies [0.04-0.15] Hz, HF: power in the high frequencies [0.15-0.4] Hz.

and construct a linear combination orthogonal to the distortions. In this algorithm, referred to as CHROM, the iPPG signal, S , is defined as:

$$S = X_s - \alpha Y_s \quad (7.2)$$

with $\alpha = \frac{\sigma(X_s)}{\sigma(Y_s)}$ and X_s and Y_s the color difference signals defined as :

$$\begin{aligned} X_s &= 0.77R_n - 0.51G_n \\ Y_s &= 0.77R_n + 0.51G_n - 0.77B_n \end{aligned} \quad (7.3)$$

the normalized color channels (R_n , G_n , and B_n) are defined as:

$$R_n = \frac{R_s}{0.7682}, G_n = \frac{G_s}{0.5121}, B_n = \frac{B_s}{0.3841} \quad (7.4)$$

with R_s , G_s , and B_s , the average pixel intensity values at frame n for the corresponding channel. Better results were obtained when band-pass filtered versions of R_n , G_n and B_n were used. This algorithm showed superior performance compared to BSS-based methods for videos with subjects performing physical exercise, including periodic motion. In addition, overlap-add intervals of only 1.6 s were used, while processing intervals of 25 s were required to give satisfactory results with the BSS-based methods.

The same authors published another algorithm, in which the blood-volume pulse signature was used to improve robustness against motion [104]. This approach did not require any assumption about the periodicity of the signals and distortions. The normalized blood-volume pulse vector, P_{bv} , defined as the signature of blood volume change, was estimated from the relative pulsilities of the color channels and used to discriminate the distortions from the pulse signal. This algorithm, referred to as PBV, showed a better robustness against motion compared to the BSS-based and CHROM methods. However, the performance of PBV was slightly inferior than that of CHROM for stationary subjects. The length of the overlap-add window was also investigated and the best results were obtained for a window duration of approximately three seconds. The authors also proposed to use the PBV and CHRO methods for the component selection in BSS-based methods, resulting in two additional methods: principal component analysis (PCA) *guided by* CHROM and PCA *guided by* PBV, which produced good results.

Later, Wang *et al.* introduced another model-based iPPG signal derivation technique called POS (plane orthogonal to skin) in [96]. The main idea behind this algorithm is to project the signal onto a plane orthogonal to the temporally normalized skin tone, in order to improve the separation between pulsatile and specular components. The following assumption was made in order to define the projection-axes:

$$\text{skin pulsatility green} > \text{skin pulsatility blue} > \text{skin pulsatility red} \quad (7.5)$$

The different steps to derive the iPPG signals using the POS algorithm are summarized in Figure 7.4. In order to assess the performance of the POS method, the authors computed an SNR measure, which was also reported for the iPPG signals computed with different existing iPPG signal derivation techniques (averaging, PCA, ICA, CHROM, PBV, SSR). The database consisted of a total of 60 video sequences, divided in four subsets with various experimental conditions (various skin tones, different luminance conditions, recovery after exercise and fitness exercise). Promising results were obtained with the POS method, which was almost always ranked as best or second-best according to the SNR metric used.

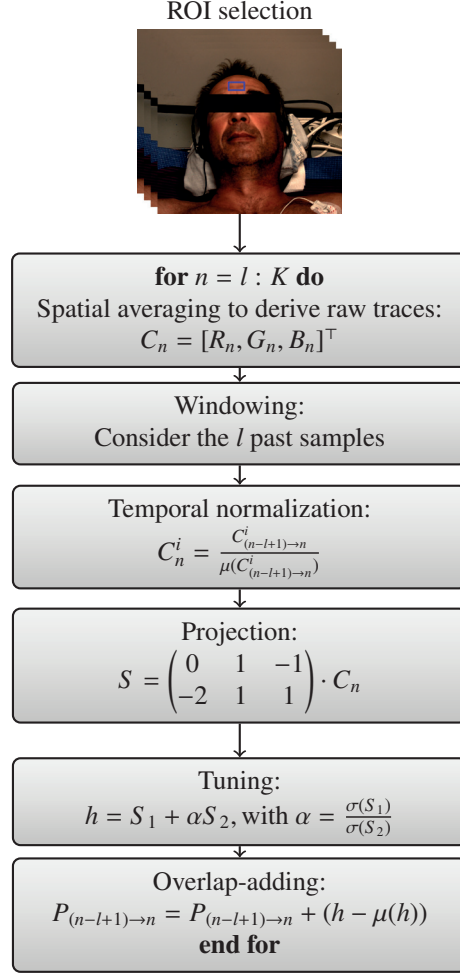


Figure 7.4: The different steps involved in the POS method [96]. C_n : vector containing the raw R/G/B traces at frame n , S : projected signal, h : short-term pulse signal, P : long-term pulse signal, K : number of frames.

In [105], an optical iPPG signal model was proposed and used to derive a processing approach aiming at attenuating motion artifacts. According to this optical model, the raw iPPG signal was represented as:

$$I_i(t) = \alpha_i \beta_i (S_0 + \gamma_i S_0 \text{Pulse}(t) + R_0) M(t), \text{ with } i \in \{R, G, B\} \quad (7.6)$$

where α_i is the power of the i th color channel in the normalized practical illumination spectrum, β_i is the power of the i th color light in the normalized diffuse reflexion spectrum, S_0 is the average scattered light, γ_i is the (ac/dc) ratio of the i th color light, $\text{Pulse}(t)$ is the normalized ideal iPPG signal, R_0 is the diffuse reflection light, and $M(t)$ is the motion modulation of the ideal iPPG signal. A color difference method was used to attenuate motion artifacts. This operation consists of a weighted subtraction between two iPPG channels. In that case, the color difference signal, $D(t)$, was written as:

$$D(t) = \frac{I_i(t)}{\alpha_i \beta_i} - \frac{I_j(t)}{\alpha_j \beta_j} \quad (7.7)$$

Similarly to the work presented in [96], the authors assumed that skin pulsatility green > skin pulsatility blue > skin pulsatility red. For this reason, the color difference operation was performed between the green and red channels, resulting in an adaptive green-red difference operation. The products $\alpha_i\beta_i$ for the green and red channels were approximated by:

$$\tilde{\alpha}_G(t)\tilde{\beta}_G(t) = \frac{I_G(t)}{\sqrt{I_R^2(t) + I_G^2(t) + I_B^2(t)}}, \quad (7.8)$$

and

$$\tilde{\alpha}_R(t)\tilde{\beta}_R(t) = \frac{I_R(t)}{\sqrt{I_R^2(t) + I_G^2(t) + I_B^2(t)}}, \quad (7.9)$$

and the green-red difference (GRD) signal was defined as:

$$GRD(t) = \frac{I_{Gf}(t)}{\tilde{\alpha}_G(t)\tilde{\beta}_G(t)} - \frac{I_{Rf}(t)}{\tilde{\alpha}_R(t)\tilde{\beta}_R(t)} \quad (7.10)$$

with $I_{Gf}(t)$ and $I_{Rf}(t)$, the band-pass filtered $I_G(t)$ and $I_R(t)$ (0.7-4 Hz). This GRD method was compared with an ICA-based method and a CHROM-based method on subjects performing six different types of head movements. The GRD method was found to be the best performing method, for all kinds of movement types.

Spatial subspace rotation method

Instead of defining an optical model, which requires a good knowledge of the complex physical phenomena involved, one can decide to use a data-driven approach. In [106], an algorithm referred to as spatial subspace rotation (SSR) was introduced. In this algorithm, conceptually quite different from the aforementioned approaches, the temporal rotation of the RGB subspace of skin pixels is estimated to derive the pulse. Figure 7.5 summarizes the different steps involved in the SSR algorithm, namely matrix construction, eigen decomposition, projection and overlap-adding. Promising results were obtained for this method, with results comparable to those of the POS method (a performance comparison can be found in [96]). The authors pointed to the following limitations. First, a well-defined skin mask is necessary in order to avoid having non-skin clusters, which are problematic. Secondly, as for the POS method, the optimal value of the window length l depends on the subject pulse rate.

Other color spaces

The use of other color spaces for iPPG applications was investigated by Tsouri *et al.* [107]. In this study, channels from seven color spaces² were used to derive iPPG signals and compared to each other for pulse rate estimation. For each method, the HR was derived from the main frequency peak between 0.74 and 4 Hz in the PSD spectrum. Results were also compared when ICA methods were applied. Interestingly, the lowest average absolute error was obtained for the Hue channel, which performed even better than the classical RGB after applying ICA-based methods. The authors concluded that the use of color spaces derived from the RGB color space is beneficial.

2. standard RGB (sRGB), hue saturation lightness (HSL), hue saturation value (HSV), hue intensity saturation HIS, XYZ, CIE XYZ and CIE YUV.

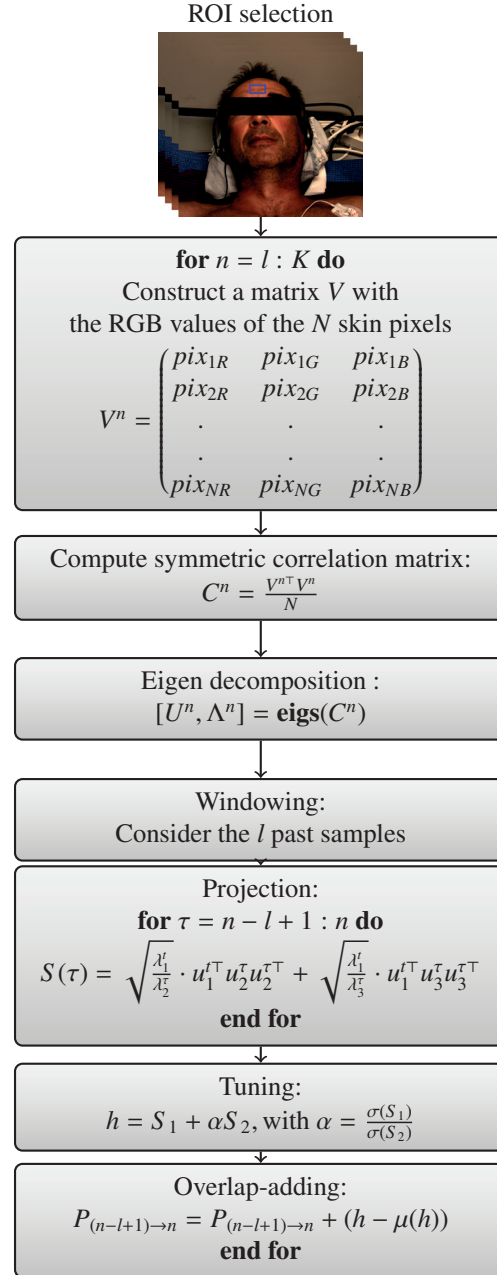


Figure 7.5: The SSR method described in [106]. V^n : matrix containing the RGB skin pixel values at frame n , C^n : symmetric correlation matrix, U^n : eigenvectors, Λ^n : eigenvalues, l : length of the sliding window, h : short-term pulse signal, P : long-term pulse signal, K : number of frames.

Other approaches

It is important to emphasize that the aforementioned list of iPPG signal derivation techniques is not exhaustive. Other methods, which will not be described in this dissertation, have been proposed. For instance, a wavelet-based filtering operation was applied for detrending and denoising

the raw iPPG signal in [108]. Another approach based on spectral graph theory was proposed in [109]. In this study, the Laplacian Eigenmap was used to map the RGB data onto a one-dimensional space, from which the blood volume pulse could be extracted. In another study, the NLMS filter was used to reduce the effects of artifacts caused by illumination variations [110]. Wu *et al.* proposed a method called Eulerian Video Magnification able to amplify subtle color variations and imperceptible motions [111]. In this approach, a spatial decomposition followed by a temporal filtering step was applied to the video frames, resulting in amplified videos. This amplification allows, for example, the visualization of blood volume flow as it fills the face at each heartbeat and could be used as a pre-processing step for iPPG applications.

Additional considerations

The selection and tracking of the ROI, which have not been discussed yet, are also very important in the extraction of iPPG signals. Selection of the ROI will be discussed in the next chapter, in which different regions on the subject face are compared. ROI tracking, for its part, is especially important when the subjects are moving. The aim here is not to provide an overview of the numerous available tracking algorithms but to keep in mind that the performance of all iPPG signal derivation methods can be enhanced by the use of a good ROI tracking algorithm. In addition, skin segmentation, i.e. automatic identification of the pixels corresponding to the skin, is also a precious tool that can improve the quality of the derived iPPG signal.

Summary

Table 7.1 summarizes the advantages and drawbacks of the iPPG signal derivation methods described in the present chapter. Several sequences recorded in the dark using a near-infrared (NIR) camera were processed in the framework of the present thesis, which has warranted mention of the applicability to an NIR setting in this table. It can be noticed that all methods except the basic averaging method require the three RGB channels and are therefore not suitable for sequences recorded with a NIR camera.

7.3 Conclusion

The contactless iPPG technology was introduced in this chapter. The principle of iPPG was first described and the most known signal-derivation methods were presented. The next two chapters focus on two applications of this technology, namely real-time HR estimation (Chapter 8) and measurement of pulse rate variability (Chapters 9). Some preliminary results for video-based HR estimation on the neonates are presented in Appendix B.

Table 7.1: Summary of the state-of-the-art iPPG signal derivation methods. BP filtering: band-pass filtering. A common choice for the pass-band is [0.7-4] Hz [112]. l : length of the sliding window [96, 106].

Method	Advantages	Drawbacks / Limitations	NIR applicable?
Averaging & BP filtering	<ul style="list-style-type: none"> • simple • small delay 	<ul style="list-style-type: none"> • not robust 	yes
BSS-based	<ul style="list-style-type: none"> • robust 	<ul style="list-style-type: none"> • component selection • large delay 	no
CHROM [103]	<ul style="list-style-type: none"> • small delay • robust to periodic motion 	<ul style="list-style-type: none"> • knowledge required to define projection plane 	no
PBV [104]	<ul style="list-style-type: none"> • small delay • robust to periodic motion 	<ul style="list-style-type: none"> • requires precise knowledge of blood volume pulse signature 	no
SSR [106]	<ul style="list-style-type: none"> • small delay • robust to motion 	<ul style="list-style-type: none"> • requires perfect skin segmentation • l parameter selection 	no
POS [96]	<ul style="list-style-type: none"> • small delay • robust to motion • built on physiological reasoning 	<ul style="list-style-type: none"> • l parameter selection 	no
GRD [105]	<ul style="list-style-type: none"> • small delay • robust to motion 	<ul style="list-style-type: none"> • requires very precise tracking 	no

A Processing Scheme for Real-Time Heart Rate Estimation

8

8.1 Motivations

There are increasing applications for contactless video-based monitoring of HR. For instance in neonatology, the attachment of sensors to the fragile skin of the neonates causes discomfort, while there is currently a lack of contactless technology. Video-based HR monitoring would therefore provide a solution to overcome the limitations of traditional sensors in such cases. Another possible application is the monitoring of drivers, in which the detected HR could be used to detect fatigue or other events that diminish driving capacity. iPPG technology could also be incorporated onto video-based surveillance systems, to monitor for example elderly people at home. The two main advantages of this technology, i.e. being contactless and relatively low-cost, make it very attractive for various applications.

The present chapter aims at introducing a novel processing scheme to estimate the HR in a contactless manner, using the iPPG signals derived from video sequences. The database as well as the acquisition system are first introduced. Preliminary tests are then performed to investigate which parts of the face result in the iPPG signals with the best quality, in order to have a well-founded basis for the ROI selection step. After these considerations for ROI selection, three different algorithms for instantaneous frequency estimation are evaluated on the derived iPPG signals. After selecting one of these algorithms, the potential benefits of combining iPPG signals derived with different methods are investigated. Finally, a challenging database subset with moving subjects is analyzed and a signal quality index is developed to increase the reliability of the HR processing scheme.

8.2 Database

The *adult database* was created at the Swiss Center for Electronics and Microtechnology (CSEM), Neuchâtel, Switzerland¹. Video sequences were recorded on 12 healthy subjects who gave their informed consent. For each subject, multiple sequences were acquired according to different experimental scenarios designed to explore different aspects of iPPG technology. Two commercial IDS® cameras with a dynamic range of 12 bits were used to record the sequences; one working in visible light² and one working in the near-infrared (NIR), for night vision³. The

1. The data acquisition was supervised by Virginie Moser and Fabian Braun from CSEM.

2. Model UI-3240CP-C-HQ with EV76C560ACT sensor.

3. Model UI-3240CP-NIR-GL with EV76C661ABT sensor.

video sequences were sample at 20 frames per second with a resolution of 1.3 megapixels. In order to assess the accuracy of the physiological parameters estimated from the video sequences, the subjects also wore classical sensors, whose signals were used as ground-truth. For this purpose, a BIOPAC® system was used to record the following waveforms: one-lead ECG, respiration (thoracic belt) and PPG (finger-transmission mode). These reference signals were synchronized with the video sequences with a precision of about one millisecond. The following experimental scenarios were used for data acquisition:

- Respiration modulation protocol: this protocol aimed at recording sequences with controlled time-varying breathing rate. Such sequences are useful to train and evaluate the respiration estimation algorithm. They are also valuable in the development of HR estimation algorithms, as modulations of breathing rate trigger HR changes. Sequences of a duration of 252 s (~4 minutes) were acquired using the protocol described in Figure 8.1.

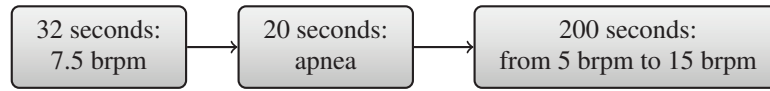


Figure 8.1: iPPG acquisition protocol: respiration modulation. brpm: breaths per minute.

- Handgrip protocol: the main purpose of the handgrip protocol was to induce HR increases without having the subjects moving. These sequences are valuable to validate the HR estimation algorithms in non-stationary conditions, which are more challenging. 4-minute sequences were acquired using the protocol described in Figure 8.2.



Figure 8.2: iPPG acquisition protocol: handgrip isometric exercise.

- Movement protocol: for this protocol, subjects were asked to rotate their head to the left and to the right according to an audio stimulus. The timings of the left and right movements of the head were randomly generated with a uniform distribution of 1-7 seconds for pauses in between the right/left transitions. After 1.5 minutes, the hand of a second person was moved in and out of the cameras field of view. For each subject, the duration of the acquired sequence was two minutes.

In addition to the different tasks performed by the subjects, different lighting conditions were tested. More specifically, the feasibility of performing video-based monitoring in the dark using a system composed of the NIR camera and an IR illumination was also investigated. The number of sequences acquired for each experimental scenario is summarized in Table 8.1. For all sequences, the subjects were lying down and the camera was placed about 1.5 meters away. The experimental setup is illustrated in Figure 8.3. It is important to mention that the subjects were wearing a white headband with a large hole opening on the forehead for reasons related to ROI tracking.

The reference HR was derived from the ECG signal. A local maxima detection was first applied in order to detect the R-waves and extract the RR intervals. The detected R-peaks were visually checked (and corrected if needed) in order to ensure the full correctness of the ground-truth. Then, the RR-intervals were uniformly re-sampled at 4 Hz to compute the true instantaneous HR. This instantaneous HR was finally averaged on 4-second windows (3-second overlap) to compute the reference HR.

Table 8.1: iPPG acquisition: The number and duration of sequences for each experimental scenarios. For the *handgrip* subset, 12 sequences were originally recorded in artificial light but a technical problem occurred for one sequence, and another was removed because the subject was moving too much.

	Artificial light RGB & NIR cameras	Darkness NIR camera
<i>respiration</i> subset	12 x 4-minute	12 x 4-minute
<i>handgrip</i> subset	10 x 4-minute	12 x 4-minute
<i>motion</i> subset	12 x 2-minute	-

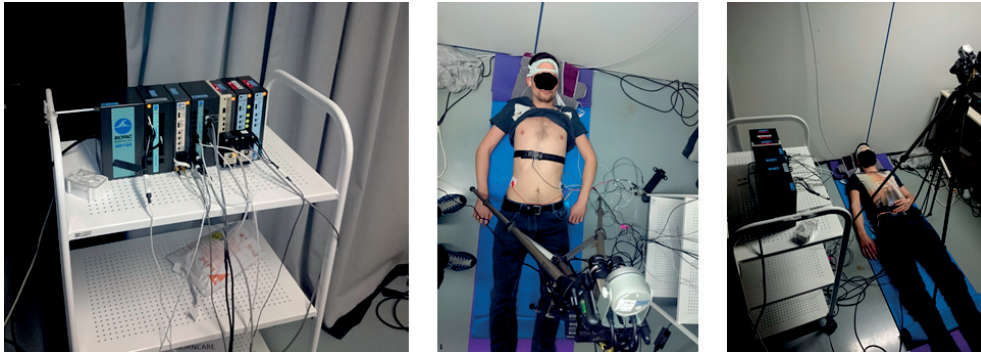


Figure 8.3: The experimental setup used for the acquisition of the *adult database*.

8.3 Methods

8.3.1 Tracking and skin segmentation

Subjects were asked to stay immobile during the recording of the *respiration* and *handgrip* subsets. ROI tracking was therefore not necessary to obtain good quality signals. However, ROI tracking was necessary for the sequences belonging to the *motion* subset. For this purpose, the tracker described in [113], based on the algorithm described in [114], was used. In this algorithm, the adaptive tracking-by-detection is based on structured output prediction achieved using a support vector machine learning framework. A budgeting mechanism is used to limit the number of support vectors and allows the tracker to run at high frame rates (20 fps in our case). In addition, a skin-segmentation step was implemented by selecting pixels whose hue and saturation match known skin color⁴. This skin-segmentation operation returns, for each frame, a binary image (called the segmentation mask) where white pixels correspond to skin and black pixels correspond to non-skin area.

The initial idea for the *motion* subset was to track a forehead ROI, in accordance with our previous findings (see Section 8.3.2). However, the fast rotations of the head present in the *motion* subset were challenging for the tracking algorithm, which had difficulties to follow the selected ROI. As a consequence, three problematic recordings of the *motion* subset were dropped. For the nine remaining recordings, some frames without skin pixels were returned during agitated periods. These frames were omitted for the computation of iPPG signals. Because of these difficulties, all the analyses of the motion subset were also performed for an ROI encompassing

4. Yann Schoenenberger and Lionnel Martin (LTS2, EPFL) developed the ROI tracking and skin segmentation methodology.

the whole face, which was easier to follow for the tracker. Examples of a tracked forehead ROI and a tracked whole face ROI are shown in Figures 8.4 and 8.5, respectively.



Figure 8.4: A sample ROI tracked on the subject forehead and the corresponding skin segmentation mask, where white pixels correspond to pixels automatically detected as skin.



Figure 8.5: A sample ROI encompassing the whole face and the corresponding skin segmentation mask, where white pixels correspond to pixels automatically detected as skin.

8.3.2 Determination of the best region on the face to derive iPPG signals

In the literature, the skin ROI chosen to compute the iPPG signals is most often selected on the subject's face. For instance, it has been shown that an ROI encompassing the whole face area can be used to recover the iPPG signals and further estimate HR [100, 115]. However, eye movements and blinking can induce artifacts. For this reason, McDuff *et al.* decided to exclude the region around the eyes from the whole-face ROI [102]. The selection of the ROI was also questioned in a study by Lewandowska *et al.*, in which two ROIs of different sizes were compared [101]. Similar accuracies were reported for an ROI on the forehead and a whole-face ROI, and the authors concluded that the selected forehead ROI was comparable to the whole-face region. Additional investigations were carried out by Verkruysse *et al.*, in a study with the iPPG signals derived from different ROIs being visually compared [5]. More specifically, four different ROIs were considered, namely the whole face, the upper head, a forehead rectangle and a single pixel on the forehead. The authors concluded that the selection of ROI is not critical for the determination of HR, but they stressed the fact the SNR can be reduced by averaging a larger number of pixels. On the other hand, when the whole face is used, some pixels do not contribute to HR signal, decreasing the signal quality. Consequently, for iPPG applications, a compromise has to be made between the number of pixels in the ROI and their relevance.

The aforementioned studies have shown promising results for HR estimation with the iPPG contactless technology. However, few researchers have addressed the problem of ROI selection and

we noted that the reasons behind the choice of the ROI remain under-studied. This study aims at providing some rationale for the choice of the ROI in iPPG applications. The results were published in a conference paper [116].

Methods

The small database used for this analysis consists of six sequences that were recorded using the same experimental setup as for the *adult database* described in Section 8.2, except that the entire face of the three subjects was visible (no headband). Images were recorded with the two cameras (RGB and NIR) simultaneously. For each video sequence, a rectangular box encompassing the whole face of the subject was manually selected. Then, this area was divided into 260 (20x13) small ROIs of equal size (see Figure 8.6 (a)). For each ROI, the raw iPPG signals were computed by spatial averaging of the pixels, for each channel and at each frame. As a pre-processing step, the resulting signals were band-pass filtered between 0.6 and 4 Hz using a Butterworth band-pass filter (forward and backward filtering). Then, a PSD analysis was performed in order to compare the percentage of power in the HR band for the different ROIs. More specifically, a 10-second centered sliding window (50% overlap) was used to compute the PSD of the iPPG signals, using an autoregressive-based parametric spectral estimation algorithm. A model order of 30 was empirically chosen. For each window and each channel, the percentage of power within a frequency band of 0.2 Hz centered at the local true HR was computed. The local true HR was defined as the average true HR derived from the ECG in the corresponding 10-second window. The resulting power percentages were averaged over the duration of the sequence. In order to facilitate the interpretation of the quantitative results, some ROIs were clustered according to the segmentation mask shown in Figure 8.6 (b) and their results were averaged. The following four sub-regions were considered: the forehead, the cheeks, the whole face and the whole face except forehead and cheeks.

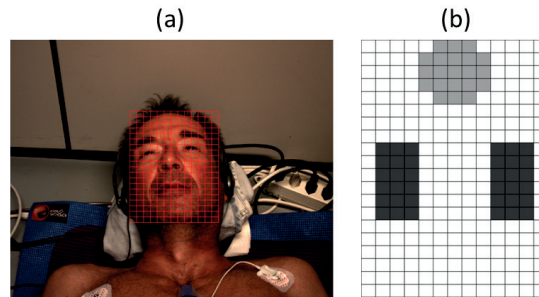


Figure 8.6: Division of the face area into 260 small ROIs which were individually processed. (a) For each subject, the grid size was fitted to the subject face size. (b) Segmentation mask. The lighter part corresponds the forehead area, and the darker parts correspond to the cheeks.

Results

Figure 8.7 shows the power percentages, averaged over all sequences, for the different clusters of ROIs representing the face regions and for the different channels. Table 8.2 reports, for each sequence, the average percentage of power in the HR band, for the different clusters of ROIs representing the face regions and for the different channels. In order to have a visual assessment of the results, power values for the different ROIs were plotted in color maps. The normalized color maps averaged over all subjects are shown in Figure 8.8.

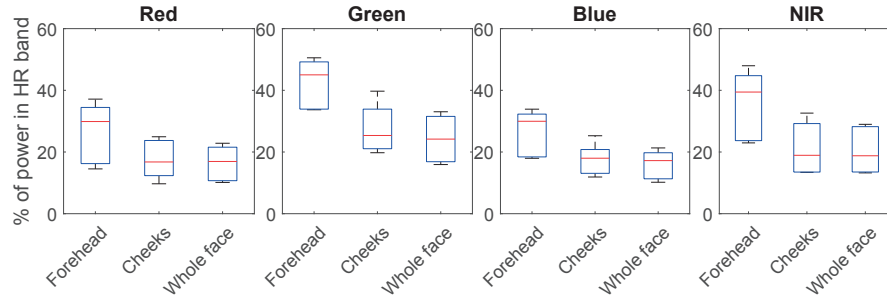


Figure 8.7: Percentages of power at true HR for the different face regions and channels.

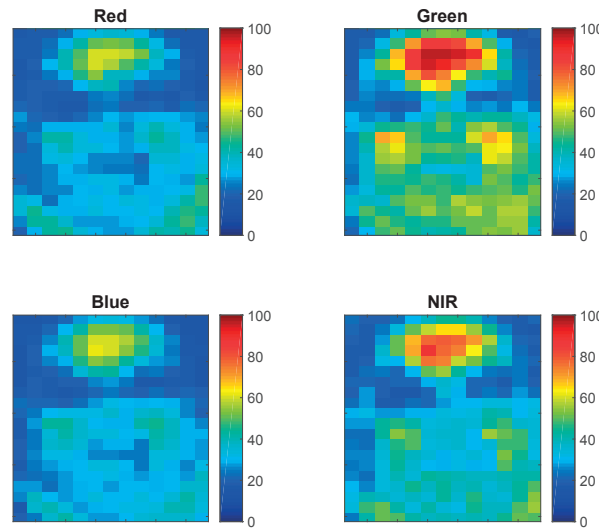


Figure 8.8: Normalized color maps averaged over all subjects to represent the amount of power at the true HR.

Discussion and conclusion

Figure 8.7 and Table 8.2 indicate that, for all channels, the forehead region was the one that contained the largest power at the HR frequency. This effect was observed for all sequences and for all channels. This is also confirmed by the average color maps (see Figure 8.8), in which the superiority of the forehead is clearly distinguishable. The color maps also indicate that cheekbones may be suitable ROIs. Interestingly, it can be seen that there was some power at HR on the neck (visible because the subjects were lying down). This is not really surprising as a lot of blood vessels pass through the neck. Concerning the cheeks regions, the power at HR frequency was larger than for the whole face 83% of the time. However, these differences were not very large. On the other hand, the difference between the whole face region and the forehead region was very marked. Indeed, the whole face region contains, on average, only 59% of the forehead information. It should also be noted from Table 8.2 that the absolute power percentage results are very variable between the three subjects, despite similar experimental conditions. This suggests that characteristics such as skin color, skin thickness and blood perfusion probably play

Table 8.2: Average percentage of power in the HR band for the different regions: (a) forehead; (b) cheeks; (c) the whole face; (d) the whole face except the forehead and the cheeks.

		Red				Green			
Sequence #		(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
subject1 {	1	32.5	24.9	22.8	21.5	48.7	39.7	33.0	30.2
	2	37.1	23.7	21.6	19.6	50.5	33.9	31.5	29.2
subject2 {	3	14.5	9.7	10.7	10.5	33.7	19.8	16.9	14.6
	4	16.2	12.3	10.1	9.1	33.9	21.5	15.9	13.1
subject3 {	5	27.3	13.1	14.4	13.3	41.3	21.1	20.9	18.8
	6	34.4	20.4	19.5	17.7	49.2	29.2	27.4	24.8
Av.		27.0	17.4	16.5	15.3	42.9	27.5	24.3	21.8

		Blue				NIR			
Sequence #		(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
subject1 {	1	32.3	25.3	21.3	19.5	44.2	32.6	28.2	25.8
	2	31.7	20.8	18.9	17.2	48.0	29.2	29.0	27.0
subject2 {	3	18.0	11.9	11.3	10.5	23.7	13.4	13.2	12.1
	4	18.4	13.1	10.2	8.8	23.0	15.6	13.5	12.2
subject3 {	5	28.2	15.3	15.5	14.3	34.6	13.5	15.6	14.0
	6	33.9	20.7	19.8	18.1	44.7	22.3	22.0	19.5
Av.		27.1	17.8	16.2	14.7	36.4	21.1	20.3	18.4

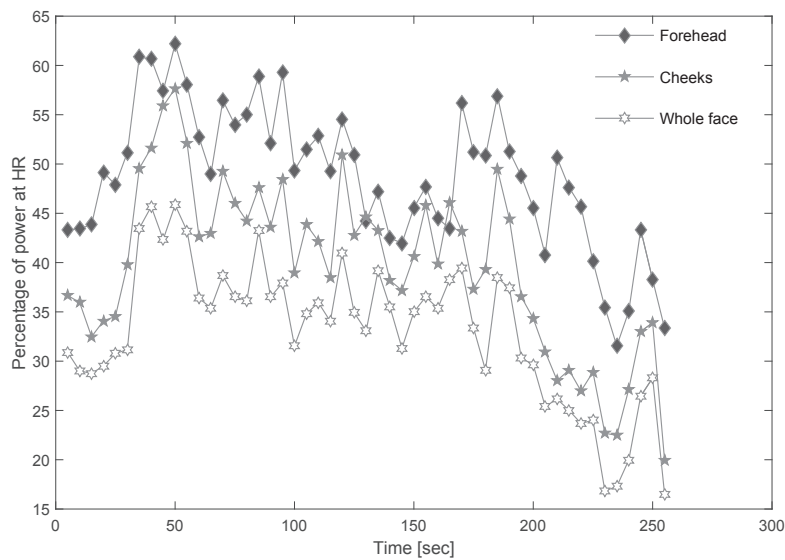


Figure 8.9: Time evolution of the power at HR frequency in the different zones for the first sequence.

an important role in the quality of the extracted iPPG signals.

Concerning the variability between the different channels, our results show that the green channel has the strongest iPPG signals (see Figure 8.7). These results are in accordance with previous findings [5, 96, 105], and may be explained by the position of the hemoglobin absorption peaks, as discussed in Chapter 3 in the context of PPG.

Figure 8.9 illustrates the time evolution of the power percentage in different face regions for the first sequence. It can be noted that the three curves follow the same pattern, indicating that it would be of a dubious interest to use the different regions in a complementary way.

It should be mentioned that the main limitation of this preliminary study is the size of the dataset. Regardless, similar tendencies were observed for all sequences and all subjects, reinforcing the hypothesis that the results are reproducible. The obtained results call into question the choice of the ROI. Indeed, most often the whole face is used [100, 115], but better performance might be achieved using only the forehead. A previous study pointed out considerations about the ROI size and its impact on the SNR [5]. Our findings highlighted the fact that, in addition to the size, the location of the ROI also matters. Indeed, the selected ROIs of equal sizes led to different results, even the ones containing only skin.

To conclude, these preliminary results suggest that the forehead part is the most appropriate region to estimate HR, followed by the cheeks. Similar conclusions were reported in other studies [117, 118], in which different ROIs were compared. Therefore, these regions should be tracked in priority in iPPG applications. Further work should be performed to validate these conclusions on a broader population, including subjects with dark skin and for different subject postures.

8.3.3 Evaluation of three algorithms to compute instantaneous HR from the iPPG signal

After determining what parts of the face contain the strongest cardiac pulsatility, investigations were pursued to find an algorithm suitable to estimate HR from the iPPG signal. Different approaches exist to derive the HR from the iPPG signal. In many studies, the HR was estimated from the spectrum of the iPPG signal, usually computed using the fast Fourier transform (FFT) algorithm [5, 99, 101, 103–105, 110, 119–123]. In other studies, time-domain methods were used to detect individual pulses and derived HR [100, 102, 109, 124]. In [125], the poles corresponding to the HR on an auto-regressive model were identified. Although promising results were obtained in the aforementioned studies, it should be noted that most of proposed approaches are based on block-wise implementations with considerable processing delays (from 7.5 seconds to one minute). For medical applications, such as monitoring in the ICU, the estimation delay should be as short as possible.

The present study aims at reporting the performance of three candidate algorithms potentially suitable for real-time HR monitoring using the iPPG contactless technology. These algorithms are: 1) an algorithm based on adaptive sliding-window singular value decomposition (SWASVD), 2) an adaptive band-pass filter (OSC-MSE-W), 3) a notch-filter bank (NFB) estimation method. In order to assess the performance of these algorithms in presence of large HR fluctuations, the *handgrip* and *respiration* datasets described in Section 8.2 were used. In addition, the possibility of performing HR monitoring in the dark using the NIR camera and appropriate illumination was also investigated. The results were published in a conference paper [126].

Methods

In the present study, the averaging iPPG signal derivation technique was used. The raw iPPG traces were obtained by averaging the pixels within a manually determined and fixed rectangular ROI on the forehead. These raw traces were then band-pass filtered between 0.6 and 4 Hz using

an 8th order Butterworth band-pass filter. Only the green channel was considered for the sequences recorded with the RGB camera, resulting in an $iPPG_{\text{green}}$ signal for sequences recorded with the RGB camera and an $iPPG_{\text{NIR}}$ signal for sequences recorded in the dark. It should be noted that filtering using an IIR filter induces signal distortions, as some frequency components are more delayed than others. This can be avoided by zero-phase filtering, i.e. filtering the signal in the forward and backward directions, which can not be applied in real-time applications. An alternative would have been the use of a FIR filter. However, it has been shown that, for similar performance, IIR filters have a dramatically lower order [127]. As the aim here was not to study the pulse shape, but to extract HR from the main oscillatory component with the smallest possible delay, an IIR filter was selected.

SWASVD-based algorithm: Singular value decomposition (SVD) is a very popular tool for subspace estimation⁵. However, SVD algorithm are generally computationally demanding and are therefore not optimal for applications that require fast tracking techniques. Recently, efforts have been put to develop adaptive SVD-based techniques with lower computational cost that can be employed for tracking applications. The sliding-window adaptive singular value decomposition (SWASVD) algorithm described in [128] was implemented. This algorithm is derived from the sequential bi-iteration SVD approach [129] and yields a recursive computation of the subspace components in a sliding window. As shown in Figure 8.10, in this algorithm the input data are first projected onto the first decomposition of the previous step (Q_A). A QR decomposition⁶ is then performed, yielding to another basis Q_B . The projection of the input data onto this basis yields the basis for the next step. It was shown that the columns of Q_B converge to the r dominant left singular vectors, while the columns of Q_A converge to the r right singular vectors and R_B and R_A converge to Σ [129].

In the present study, a 50-sample causal sliding window (step size of one sample) and six subspaces were used. At each time step, the frequency of each subspace component ($f_1(t), \dots, f_6(t)$) was computed from the discrete Fourier transform (DFT) of the matrix with columns converging to the right dominant singular vectors (Q_A). In order to enhance the precision of the frequency estimate, a two-step procedure was used. First, the frequency bin corresponding to the largest DFT sample was extracted. Then, an iterative frequency estimation was performed, as proposed in [130]. Our observations showed that, most of the time, the oscillation of interest (i.e. reflecting the blood volume changes) was contained in the first component of the decomposition. However, in some situations (artifacts, bad-quality waveforms), it corresponded to the second component. For this reason a supplementary mechanism was developed to automatically switch to the second component f_2 in such cases. More specifically, at each time step the median of f_1 for the previous 250 samples was computed. Then, if the value $f_2(t)$ was closer to this median value compared to $f_1(t)$, the algorithm automatically switched to $f_2(t)$. Figure 8.11 summarizes the successive processing steps involved in the SWASVD-based algorithm.

OSC-MSE-W algorithm: The OSC-MSE-W algorithm described in Section 2.3.1 was used. In this study, the parameter reflecting filter bandwidth (β) and the two forgetting factors (δ and μ), were set to 0.97. The $iPPG_{\text{green}}/iPPG_{\text{NIR}}$ and smoothed versions of this signal (causal window, filter lengths of 3 and 5 samples) were provided as inputs for the light/dark conditions, respectively.

NFB algorithm: This algorithm uses a bank of notch filters to measure the common instantaneous frequency in several input signals [131]⁷. More specifically, the filter bank is composed of

5. The SVD of X is the factorization $X = U\Sigma V^H$, where U and V are orthonormal matrices and Σ , a rectangular diagonal matrix. The diagonal entries σ_i are known as the singular values.

6. A QR decomposition of a matrix A is a composition of this matrix into a product such as $A = QR$, where Q is an orthogonal matrix and R is an upper triangular matrix.

7. Developed by Leila Mirmohamadsadeghi, former PhD student of our group.

Initialize : $\mathbf{Q}_A(0) = \begin{bmatrix} \mathbf{I}_r \\ \mathbf{0} \end{bmatrix}$; $\mathbf{Q}_B(0) = \begin{bmatrix} \mathbf{I}_r \\ \mathbf{0} \end{bmatrix}$; $\mathbf{R}_A(0) = \mathbf{I}_r$;		
FOR EACH TIME STEP DO :		
Input : $\mathbf{x}(t)$		
First Iteration :		Complexity :
$\mathbf{h}(t) = \mathbf{Q}_A(t-1)^H \mathbf{x}(t)$		$8Nr$
$\begin{bmatrix} \mathbf{B}(t) \\ \times \dots \times \end{bmatrix} = \begin{bmatrix} \mathbf{h}(t)^H \\ \mathbf{Q}_B(t-1) \mathbf{R}_A(t-1)^H \end{bmatrix}$		$4Lr^2$
$\mathbf{B}(t) = \mathbf{Q}_B(t) \mathbf{R}_B(t)$		$19Lr^2$
Second Iteration :		
$\mathbf{x}_\perp(t) = \mathbf{x}(t) - \mathbf{Q}_A(t-1) \mathbf{h}(t)$		$8Nr$
$\mathbf{A}(t) = \mathbf{Q}_A(t-1) \mathbf{R}_B(t)^H + \mathbf{x}_\perp(t) \mathbf{q}_{B_1}(t)^H$		$4Nr^2$
$\mathbf{A}(t) = \mathbf{Q}_A(t) \mathbf{R}_A(t)$		$19Nr^2$

Figure 8.10: SWASVD algorithm. Image from [128], © 2004 IEEE.

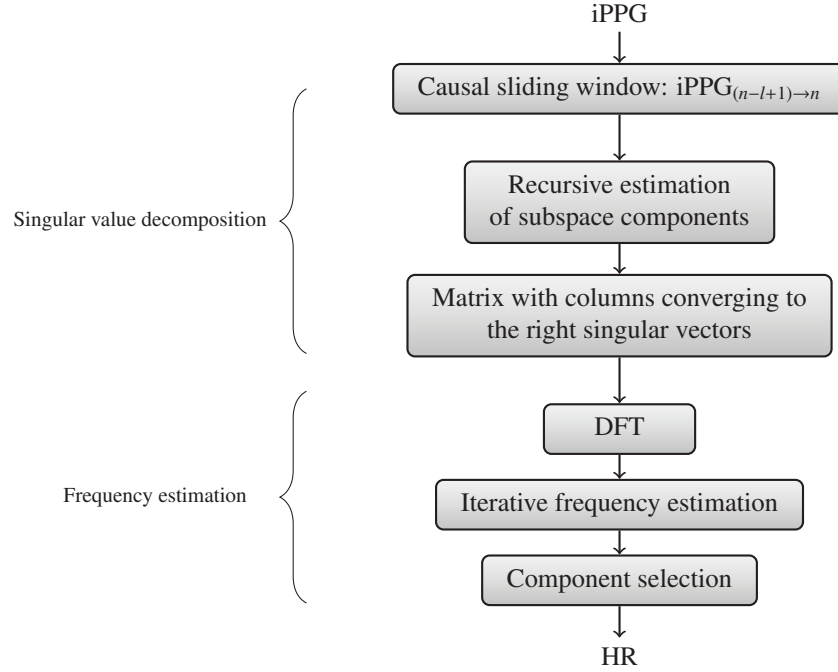


Figure 8.11: SWASVD-based scheme for HR estimation.

length-3 FIR notch filters, equally spread over a range of discrete frequencies. In this setting, the main frequency of the signal is closer to the notch frequency of the filter resulting in the smallest output. The output-to-input power is computed and then used to calculate a set of weights. The frequency estimate is finally derived from the weighted sum of the notch frequencies. The multiple inputs are combined in a similar scheme to the one used in the OSC-MSE-W algorithm. In this study, the $iPPG_{\text{green}}$ / $iPPG_{\text{NIR}}$ and smoothed versions (causal window, filter lengths of 3 and 5 samples) were provided as inputs for the light/dark conditions, respectively. The number of filters in the filter bank was chosen to be 50 in the frequency band [0.5-5] Hz. Figure 8.12 summarizes the successive processing steps of the NFB algorithm.

Performance assessment: The average absolute error (AAE) was computed between the

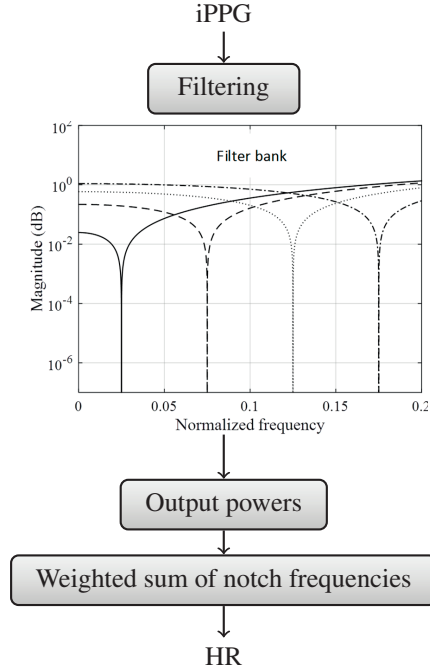


Figure 8.12: The different steps of the NFB algorithm.

true and the estimated HR values. For this purpose, HR estimates were averaged on 4-second windows (3-second overlap). A Bland-Altman analysis was also performed to define another measure of the similarity between the true and the estimated HR values. The limits of agreement (LOA) $[\mu - 1.96\sigma, \mu + 1.96\sigma]$ to contain 95% of the differences were calculated. In addition, the estimation delay was computed using the cross-covariance between the true and the estimated HR. The average delay of each algorithm was then used to align the HR sequences and to re-compute the previously described performance metrics for the aligned HR sequences and provide algorithm accuracy independently of its delay. Occasionally, artifacts on iPPG waveforms were induced by small movements of the subjects. These artifacts were characterized by sudden amplitude increases in the iPPG waveforms. In order to avoid large estimation errors, a method based on the amplitude ratio of consecutive signal windows was developed to automatically detect and remove these segments. A more elaborated quality measure for iPPG signals is presented in Section 8.3.5.

Results

Figure 8.13 shows a segment of iPPG signal derived from a video sequence acquired using the NIR camera in total darkness. The smoothed iPPG signals, which were also provided as inputs to the OSC-MSE-W and NFB algorithms, are also displayed.

In order to quantify the HR fluctuations induced by the respiration and handgrip tasks, the HR range was computed for each record as $(HR_{max} - HR_{min})$. Among all records, the average range was 25.4 ± 8.1 bpm. Figure 8.14 shows an example of the estimated HR for a subject performing the handgrip exercise. An averaged estimation delay of four seconds was found for the SWASVD and OSC-MSE-W algorithms and of three seconds for the NFB algorithm. The AAE values of each sequence have been plotted in Figure 8.15, for both visible light and dark conditions (delay compensated). The results for the overall AAE and the LOA for sequences

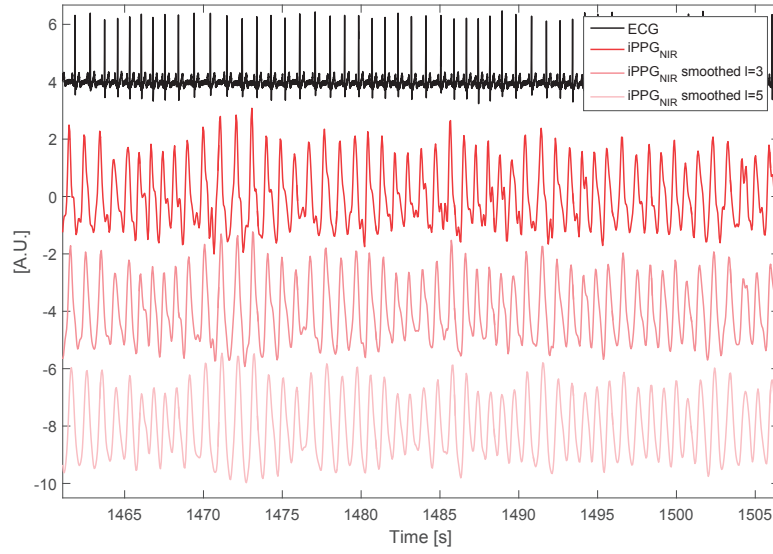


Figure 8.13: Sample of $iPPG_{NIR}$ signal derived from a sequence acquired in total darkness.

recorded in artificial light and in the dark are shown in Tables 8.3 and 8.4, respectively. These performance metrics are reported for both the real-time setting and for the aligned HR series. In total, 1.7% of HR estimates were removed based on the detection of bad-quality segments.

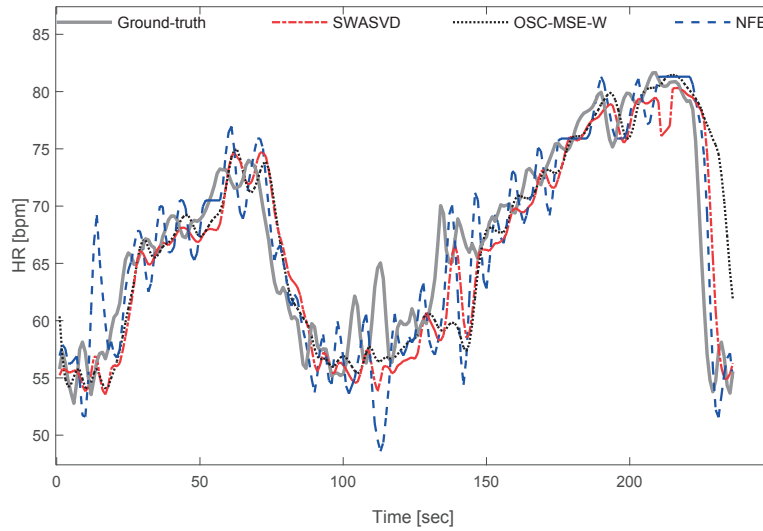


Figure 8.14: Estimated HR values and ground-truth for a recording of the *handgrip* subset, in artificial light.

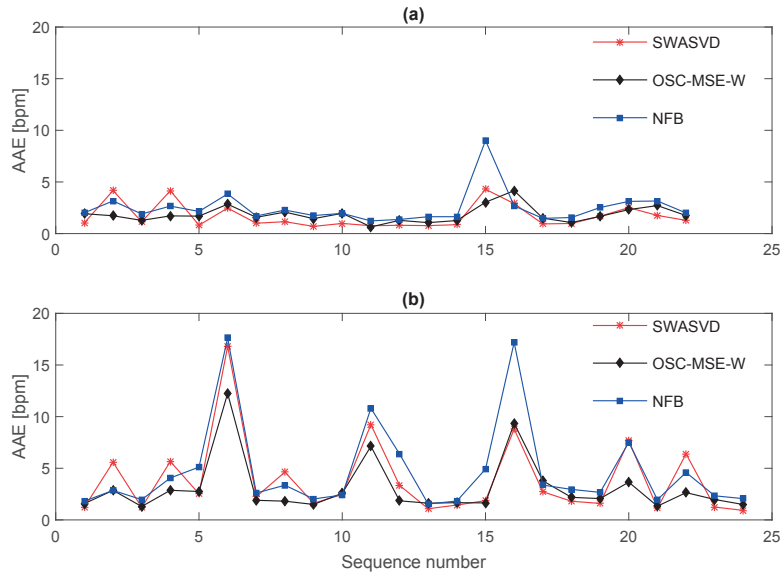


Figure 8.15: AAE for HR estimation from iPPG for all sequences recorded in visible light (a), and in the dark (b). The estimation delay was compensated.

Table 8.3: AAE and LOA results for the three algorithms and for the sequences recorded in visible light.

		SWASVD	OSC-MSE-W	NFB
Real-time	mean AAE [bpm]	3.42 ± 1.17	3.14 ± 0.84	3.98 ± 1.49
	LOA 95% [bpm]	[-8.73 10.60]	[-7.79 9.65]	[-11.03 12.60]
Aligned *	mean AAE [bpm]	1.69 ± 1.19	1.85 ± 0.78	2.49 ± 1.61
	LOA 95% [bpm]	[-6.99 8.90]	[-6.59 8.50]	[-9.26 10.87]

* Aligned means that the estimation delay of each algorithm was compensated.

Table 8.4: AAE and LOA results for the three algorithms and for the sequences recorded in the dark.

		SWASVD	OSC-MSE-W	NFB
Real-time	mean AAE [bpm]	5.25 ± 3.43	4.21 ± 2.51	6.02 ± 4.00
	LOA 95% [bpm]	[-12.67 18.46]	[-10.66 13.77]	[-17.05 20.97]
Aligned *	mean AAE [bpm]	3.88 ± 3.76	3.07 ± 2.71	4.75 ± 4.46
	LOA 95% [bpm]	[-11.76 17.59]	[-9.51 12.66]	[-15.98 19.94]

* Aligned means that the estimation delay of each algorithm was compensated.

Discussion

Overall results are encouraging and confirm that all the investigated algorithms were suitable to estimate the HR from video-sequences of the subjects. In order to better emphasize the accuracy, and given that large HR variations were present in the database, it was important to report the performance also when the estimation delay was compensated. Although better results were obtained for the video-sequences recorded in visible light, HR monitoring in the dark using an NIR camera and infrared illumination seems conceivable with an average AAE under 5 bpm (see Table 8.4 and Figure 8.15). The results are in the same range for the three algorithms. In terms of AAE, the best average performance was achieved by the OSC-MSE-W algorithm. More precisely, this algorithm resulted in an average AAE of 3.14/4.21 bpm for visible/dark sequences with the real-time setup and 1.85/3.07 bpm when the delay was compensated. It can be noted in Figure 8.15 that the individual AAE values were very similar between the three methods (expected for some recordings that were more challenging for the NFB or SWASVD algorithms), indicating that the three algorithms were sensitive to the same kind of perturbations. The smallest estimation delay was achieved by the NFB algorithm. Concerning the choice of the input channel, the initial idea was to combine the R/G/B channels for the multiple-input algorithms (OSC-MSE-W and NFB). However, it was noticed during the development of the HR processing schemes that better results were achieved with the signals derived from the green channel only, compared to the RGB combination. This is consistent with previous findings, for which the correlation between estimated and ground-truth HR was larger for the green channel than for the RGB [102].

Variable results have been reported in prior works regarding the accuracy of HR estimation in iPPG applications. For instance, an AAE of 6.1 bpm was reported using the RADICAL ICA method, for video-sequences recorded with a webcam in stable conditions [115]. In another study, a correlation of 1 was reported between the HR derived from video-sequences and from a contact PPG sensor [102]. The pole cancellation method described in [125] resulted in an AAE of approximately 3 bpm. In [108], a wavelet-based iPPG signal derivation technique led to an average beat-to-beat RMSE of 1.97 ± 0.62 bpm on a database of 12 subjects. In another study, an FFT-based approach was used to estimate HR from an iPPG signal derived from the hue channel of the HSV color space [107]. In this study, an AAE of 4.31 ± 0.4 bpm was reported over epochs of 30 seconds. The feasibility of estimating the HR using a smartphone camera was studied in [119], where an FFT-based approach led to an average error rate of 1.08 %.

It is important to emphasize that it is debatable to compare the performance obtained in the different studies. Indeed, very variable methods and window lengths are used to compute the error. Most of the time, the error is computed on long windows (about 30 seconds) and is therefore not an instantaneous error. In addition, different types of cameras and experimental setups have been used. Still, we draw the following conclusions 1) the investigated algorithms meet state-of-the-art performance and 2) they have the shortest estimation delays. In addition, we were able to demonstrate their potential using a database containing large HR fluctuations. We believe that the presence of these large HR fluctuations gives more credibility to the results obtained.

8.3.4 Combining different iPPG signals

An important step in the development of a video-based HR estimation scheme is the selection of the technique used to derive the iPPG signals from the pixel intensity values. As mentioned in Section 7.2, different methods exist to derive the iPPG signals when the sequences are recorded using a color camera. In the previous section (Section 8.3.3), we showed that promising results were obtained for the OSC-MSE-W algorithm to perform real-time HR estimation. This algorithm can track the common oscillatory component present in multiple input signals (see Chapter 2). Hence, we decided to investigate the potential benefits of an approach combining the

iPPG signals derived using different methods. More specifically, we implemented the following methods (described in Section 7.2):

- iPPG_{green}: Spatial averaging of pixels for the green channel, and band-pass filtering between 0.6 and 4 Hz.
- iPPG_{hue}: Transformation of the RGB space to the HSV space, spatial averaging of pixels for the hue channel and band-pass filtering between 0.6 and 4 Hz.
- iPPG_{GRD}: Adaptive green-red difference [105].
- iPPG_{SSR}: Spatial subspace rotation, using a sliding window of length $l = 20$ samples (one second) [106].
- iPPG_{POS}: Plane-orthogonal-to-skin, using a sliding window of length $l = 20$ samples (one second) [96].

Figure 8.16 shows the aforementioned iPPG signals for a segment of the first sequence of the *respiration* subset of the database described in Section 8.2. Figure 8.17 also shows the iPPG signals, but for a sequence from the *motion* subset. For the *motion* subset, a tracking of the ROI was necessary. This step is detailed in Section 8.3.1. It can be noticed that the iPPG_{green} and iPPG_{hue} signals are very clean when the subject is not moving. However, the cardiac pulses are less distinguishable when the subject is moving, while they are still visible in the other waveforms (iPPG_{SSR}, iPPG_{POS}, iPPG_{GRD}).

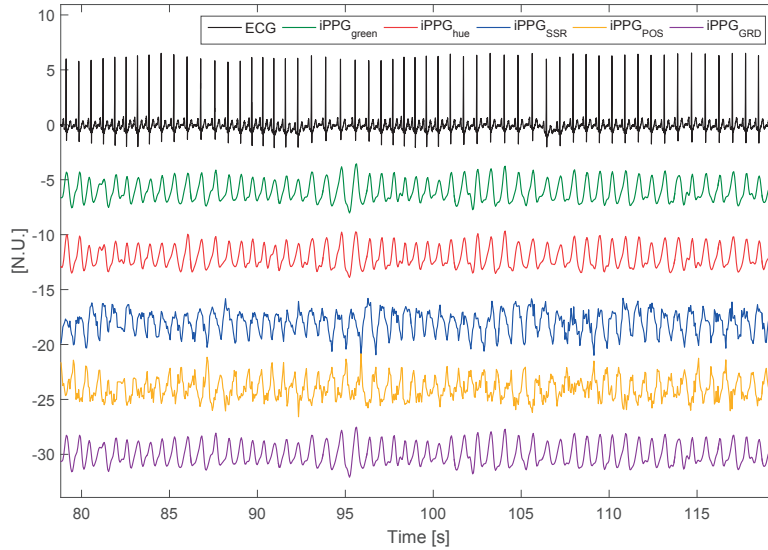


Figure 8.16: Sample of the iPPG signals derived with the different methods for a recording from the *respiration* subset. ECG is also displayed as visual reference.

In order to investigate the potential benefits of combining these different iPPG signals, the HR was estimated from each of them individually and collectively using the OSC-MSE-W algorithm. Figure 8.18 shows the individual AAE values for the *static* data subset, which includes the *respiration* and *handgrip* subsets and Figure 8.19 shows the same measure for the *motion* subset (full face tracking) (see Section 8.2). It can be noticed that 1) the signal leading to the smallest error is not always the same, 2) the performance of a given iPPG-signal derivation technique can be very variable from one sequence to another. This suggests that all methods have their advantages and drawbacks depending on the situation. Table 8.5 reports, for each iPPG signal,

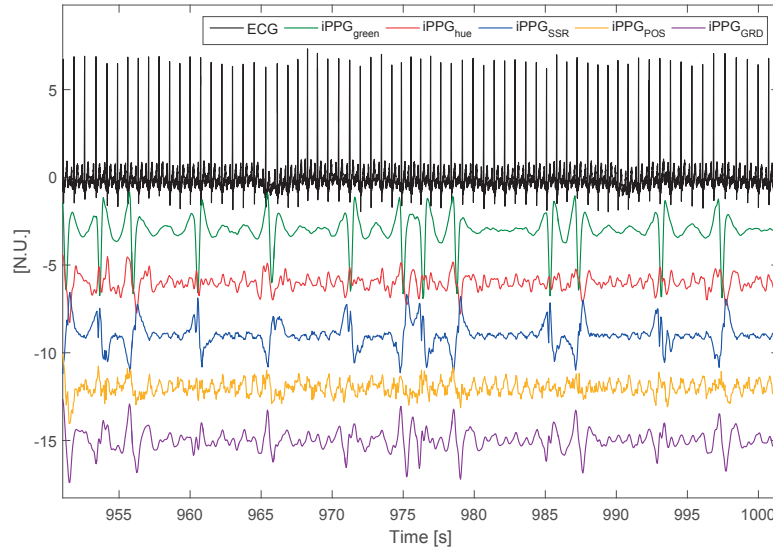


Figure 8.17: Sample of the iPPG signals derived with the different methods for a recording from the *motion* subset. ECG is also displayed as visual reference.

the number of sequences for which it led to smallest/highest AAE. It can be noted that the SSR and POS techniques seem to have a slight advantage over the other single input methods, but the combination is more advantageous. Overall, the combination led to the best results 58% of the time and never led to the worst ones. In contrast, the best single input, the $iPPG_{POS}$, led to the best results 19% of the time and to the worst ones 19% of the time. The average AAE confirms that it is more advantageous to combine the different iPPG signals. Indeed, for both datasets, the smallest average AAE was obtained for the combination of all input signals (1.40 ± 0.83 bpm for the *static* subset and 16.95 ± 16.24 bpm for the *motion* subset). Combining iPPG signals derived using different methods is therefore a good way to increase the robustness of HR estimation. It should be noted that, despite the combination of the different iPPG signals, results obtained for the *motion* subset are not satisfying for some sequences. Section 8.3.5 describes the development of a signal quality index (SQI) for iPPG signals in order to increase the accuracy of HR estimation in challenging conditions such as those of the *motion* subset.

Table 8.5: Number of times each iPPG-input resulted in the best HR estimation (# best) and in the worst HR estimation (# worst).

	<i>static</i> subset (N=22)		<i>motion</i> subset (N=9)		total (N=31)	
	# best	# worst	# best	# worst	# best	# worst
$iPPG_{green}$	0	13	2	2	2	15
$iPPG_{hue}$	0	5	0	2	0	7
$iPPG_{SSR}$	3	0	0	2	3	2
$iPPG_{POS}$	4	4	2	2	6	6
$iPPG_{GRD}$	0	0	2	1	2	1
Combination	15	0	3	0	18	0

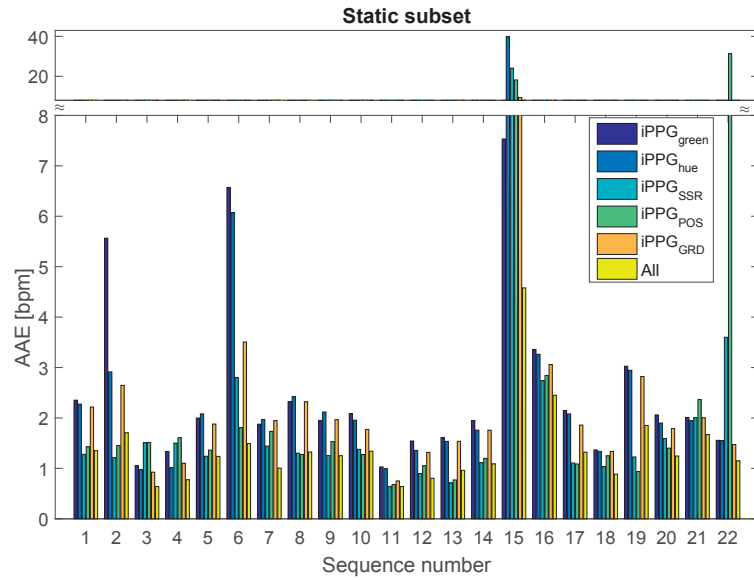


Figure 8.18: Individual AAE results for each sequence and each input signal, for the *static* subset. The estimation delay was compensated.

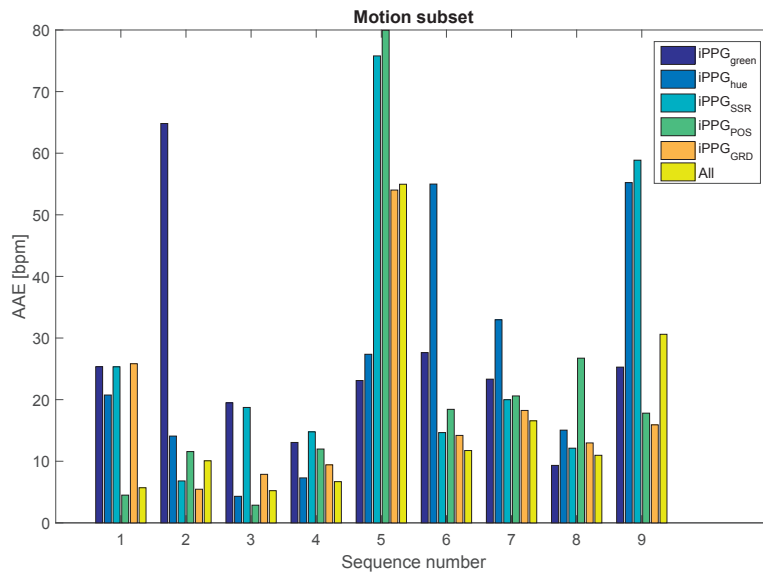


Figure 8.19: Individual AAE results for each sequence and each input signal, for the *motion* subset. The estimation delay was compensated.

8.3.5 Signal quality index

An overview of the existing iPPG-signal derivation techniques was presented in Section 7.2. However, as mentioned by Wang *et al.*, the iPPG-signal derivation technique resulting in the highest SNR depends on experimental conditions that may additionally vary over time [96]. This was confirmed by our observations, showing that the performance of the investigated iPPG signal derivation methods is quite variable from one sequence to another (see Section 8.3.4). In addition, iPPG signals are easily deteriorated by motion and illumination changes. Therefore, an SQI can be a valuable tool to increase the robustness of iPPG-based monitoring applications. More specifically, an SQI can be used to 1) detect bad-quality epochs during which the estimated HR is not reliable and 2) develop processing schemes involving a quality-based dynamic combination of iPPG signals derived with different techniques.

Two different kinds of approaches are conceivable to assess the quality of a video-derived waveform. The SQI can either be computed from features derived from pixel values of the video frames or from features derived from the iPPG signal. Both kinds of approaches have their advantages and their drawbacks. On one hand, features extracted directly from images can be obtained with a smaller delay than features extracted from the iPPG waveform. In addition, they can be related more easily to the source of the artifacts. On the other hand, these image-derived features do not depend on the technique used to derive the iPPG signals, which are not equally sensitive to specific disturbances. Therefore, features computed from the pre-processed iPPG time series have the advantage to lead to signal-dependent SQI. Being aware of these considerations, we opted for an SQI combining both kinds of information and developed a real-time signal-dependent SQI. This iPPG-SQI was the object of a conference publication [132].

Methods

The frame-to-frame average absolute difference between pairs of corresponding pixels in the ROI (DI) was first computed as follows:

$$DI[n] = \frac{1}{N_{pix}} \sum_{k=1}^{N_{pix}} |pix_k[n] - pix_k[t-1]| \quad (8.1)$$

with $pix_k[n]$, the green channel value of the k^{th} pixel of the ROI, at frame n and N_{pix} , the total number of pixels in the ROI. Sudden surges of this measure of inter-frame variability are correlated with various disturbances such as subject movements, illumination changes and camera occlusions, as shown in Figure 8.20.

When a sample is associated with a significant DI increase, empirically defined as a current DI value at least five times larger than the median DI over the past 50 samples, the sample is labeled as a candidate for the beginning of a bad-quality region (i.e. $SQI = 0$). It was observed that the iPPG signals were not all affected in the same way by the disturbances. Therefore, the final step of the SQI calculation was designed to be signal-dependent. Two configurations were considered:

- $iPPG_{green}$, $iPPG_{hue}$ and $iPPG_{GRD}$: For each labeled candidate, the SQI was set to zero for the next five seconds.
- $iPPG_{POS}$, $iPPG_{SSR}$: As shown in [96], these signals are relatively resistant against disturbances. Our observations showed that bad-quality regions were associated with local amplitude increases of these iPPG signals. For every labeled candidate, the presence of such an increase is checked using the following rule. The amplitude before the labeled candidate is first computed as: $amp_{before} = \max\{x[n-win], \dots, x[n]\} - \min\{x[n-win], \dots, x[n]\}$ with $win = 20$ samples and x , the iPPG signal of interest. After 10 samples, this amplitude is compared with the current amplitude, computed in the same way (using a causal window

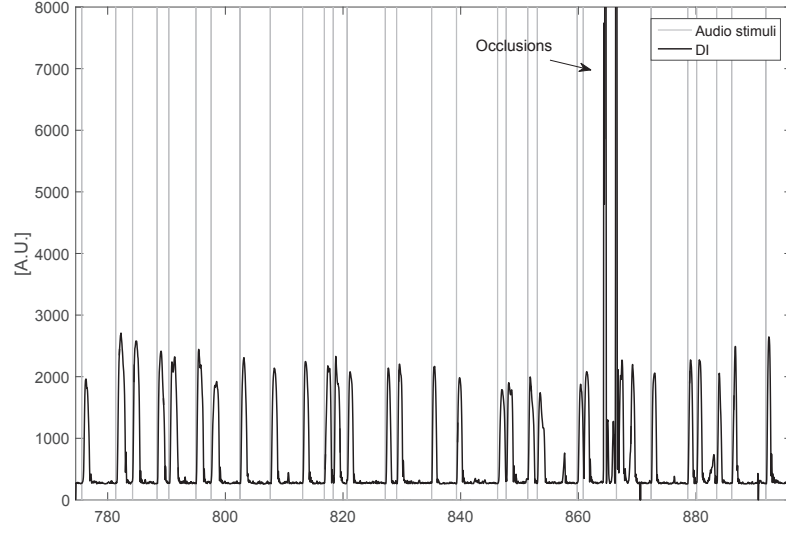


Figure 8.20: The time evolution of the DI for a sequence taken in the *motion* subset. The timings of the audio stimuli indicating to the subjects when to move are also displayed as vertical grey lines.

of 20 samples). If the ratio $amp_{current} / amp_{before}$ is larger than 1.6 (empirically selected threshold), the SQI is set to zero for the current sample. The general scheme illustrating the main steps involved in the SQI computation is shown in Figure 8.21.

The resulting SQIs were used for two purposes 1) to modify the update mechanism of the adaptive filter used to perform adaptive frequency tracking and 2) to label bad-quality segments for which the estimated HR is not reliable. HR estimation was performed using the OSC-MSE-W-SQI algorithm, which uses the input SQI(s) to freeze filter adaptation when the quality of the input signal(s) is not good (see Chapter 2, Section 2.3.3). The value of β and δ were set to 0.95. All the iPPG signals, as well as their smoothed versions (moving average using window lengths of 3 and 5 samples), were provided as inputs. It is important to stress that the frequency tracking scheme used was not affected by phase differences between the inputs

Performance assessment

In order to investigate the potential benefits of the proposed SQI, HR was estimated with and without SQI. As for the ground-truth, HR estimates were averaged on 4-second windows (3-second overlap). The accuracy of the estimated HR was quantified with the AAE, computed between the true and the estimated HR values. A Bland-Altman analysis was also performed to compute the LOA $[\mu - 1.96\sigma, \mu + 1.96\sigma]$ to contain 95% of the differences. The percentage of removed data points, considered as not trustworthy (i.e. when all the $SQI_m[n]$ were equal to zero), was also reported.

Results

Figure 8.22 is an example of the time evolution of the proposed SQI for a data segment taken from the *motion* subset, for the iPPG_{SSR} signal. The results for the performance comparison

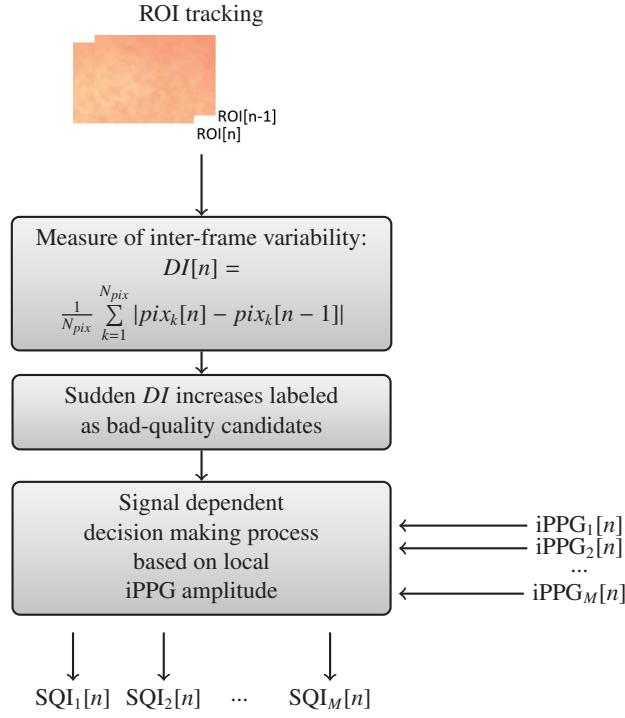


Figure 8.21: Computation of the proposed iPPG SQI. N_{pix} : number of pixels in the ROI, $pix_k[n]$: green channel value of the k^{th} pixel of the ROI at frame n .

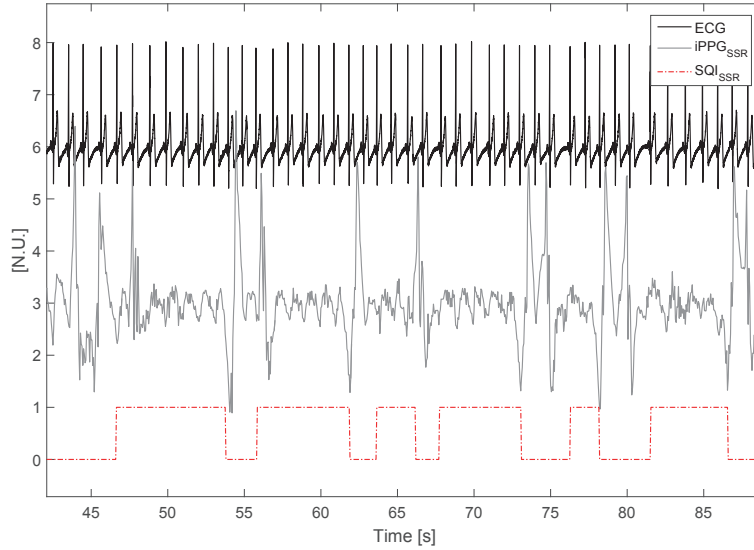


Figure 8.22: The SQI for a sequence from the *motion* subset (forehead ROI). The ECG is also displayed as visual reference

between the original scheme and the SQI-based scheme and for the two data subsets are shown in Tables 8.6 and 8.7. The performance metrics are reported for both the real-time setting (Table 8.6) and for the aligned HR series (Table 8.7), for which the averaged estimation delay of 4 seconds was compensated. The individual AAE values for the *motion* subset are plotted in Figure 8.23. Using the forehead ROI, the error was reduced by using the SQI for all the sequences and an average error reduction of 42% was achieved. According to our quality measure, a fraction of 24% of data was considered as not reliable. When the entire face was used as ROI, the error was reduced for 8 sequences by using the SQI and an average error reduction of 45% was achieved, with 13% of data considered as not reliable. An example of HR estimation can be found in Figure 8.24, where the two HR estimates, computed with the original and the SQI-based scheme, are compared. The corresponding iPPG waveforms are also shown.

Table 8.6: Performance comparison between the original scheme and the SQI-based scheme for the real-time configuration.

		Av. AAE [bpm]	LOA 95% [bpm]	% removed
<i>Static</i> subset, forehead	Original scheme	3.28 ± 0.86	[-8.62 10.05]	0
	SQI-based	3.27 ± 0.92	[-8.64 10.00]	0.01
<i>Motion</i> subset, forehead	Original scheme	16.46 ± 14.04	[-27.79 32.14]	0
	SQI-based	6.74 ± 2.77	[-16.06 19.26]	23.65
<i>Motion</i> subset, full face	Original scheme	9.89 ± 4.23	[-20.16 29.35]	0
	SQI-based	5.32 ± 4.94	[-14.31 19.51]	13.3

Table 8.7: Performance comparison between the original scheme and the SQI-based scheme for the aligned configuration.

		Av. AAE [bpm]	LOA 95% [bpm]	% removed
<i>Static</i> subset, forehead	Original scheme	1.92 ± 0.89	[-5.75 7.08]	0
	SQI-based	1.89 ± 0.93	[-5.68 6.95]	0.01
<i>Motion</i> subset, forehead	Original scheme	11.57 ± 6.21	[-24.75 28.36]	0
	SQI-based	6.21 ± 2.55	[-14.93 18.41]	23.65
<i>Motion</i> subset, full face	Original scheme	9.85 ± 4.32	[-20.41 29.70]	0
	SQI-based	4.72 ± 5.27	[-13.93 19.24]	13.3

Discussion

The accuracy of HR estimation was considerably increased by the addition of the SQI for the *motion* subset, while the results for the *static* subset remained almost unchanged, which can be explained by the small amount of subject movements in these videos. The average error of about five bpm for the *motion* subset is encouraging given the challenging nature of this dataset. Regarding the computation of the SQI, it should be mentioned that SQI metrics have been proposed already in the context of classical PPG signals. For instance, Li *et al.* proposed an SQI based on dynamic time warping for pulsatile signals (PPG and ABP) [85]. Unfortunately, it seems difficult to adapt this approach to iPPG signals, because the morphology of the latter is very variable. We believe that the proposed method based on the *DI* value has some strengths. First, it only requires the pixel values at the current frame and at the frame [n-1] and it is very

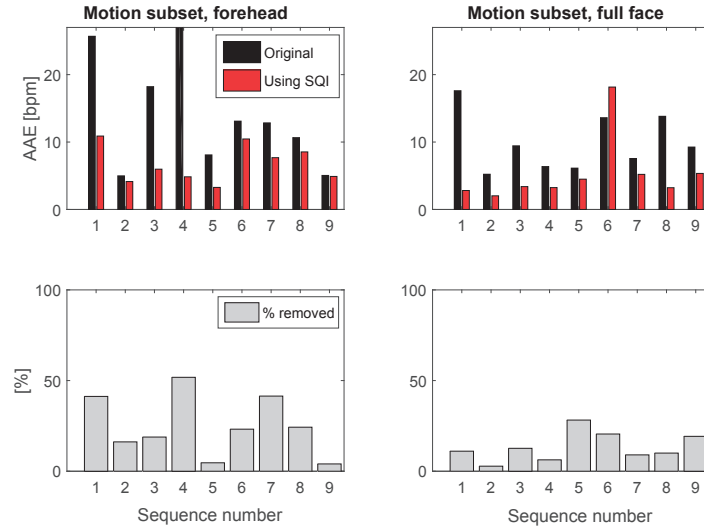


Figure 8.23: Individual AAE values and percentages of data considered as not reliable for HR estimation for the *motion* subset.

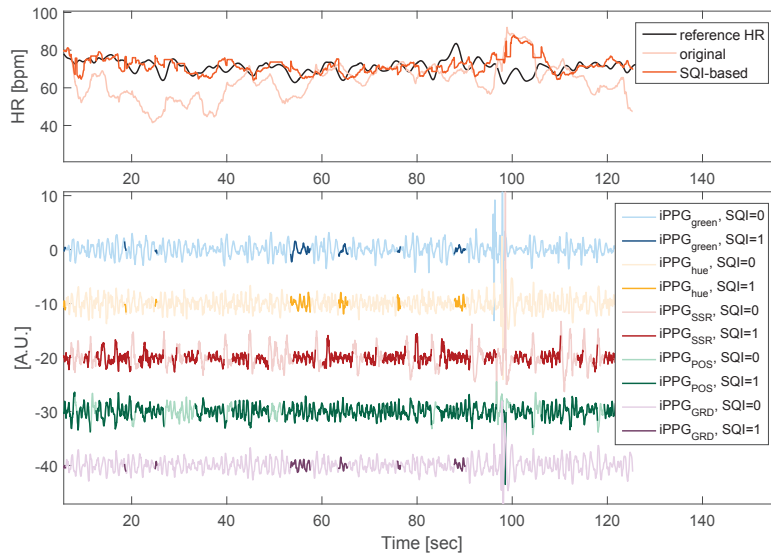


Figure 8.24: HR estimated from all iPPG signals with and without using the SQI for a sequence of the *motion* subset (top) and the corresponding iPPG waveforms, where darker segments indicate good quality segments (bottom).

simple to compute. Second, this inter-frame variability measure raises with all kinds of disturbances that are likely to deteriorate the quality of the iPPG signals. Other measures such as the time evolution of the position of the tracked ROI could have been used instead. However, in that

case, sudden illumination changes would not be detected.

In addition to the proposed SQI, we introduced in this study a complete HR estimation scheme taking into account signal quality and integrating different state-of-the-art methods to derive the iPPG signals [96, 105, 106]. As illustrated in Figure 8.24, this SQI-based adaptive frequency tracking scheme allows us to combine the relevant information present in the different iPPG waveforms, while bad-quality segments do not affect frequency estimation. In this very extreme case, it can be noticed that only iPPG_{SSR} and iPPG_{POS} contribute significantly to HR estimation. Nevertheless, it is important to keep in mind that both the frequency and the amplitude of the movements performed by the subject in the *motion* subset are not very natural.

The lack of labels for waveform quality was a limitation of this study. Future work could include testing this SQI on videos with subjects performing smoother movements, in time-varying illumination conditions.

8.4 Conclusion

The development of a processing scheme to perform real-time HR estimation from video sequences was presented in this chapter. Using a database recorded on adult subjects, different aspects of the iPPG technology were studied, leading to the following observations/findings. The best regions on the subject face to derive the iPPG signals are the forehead and the cheeks, and they lead to higher SNR compared to the entire face region for static subjects. The three studied algorithms (OSC-MSE-W, SWASVD-based and NFB) can be used to perform accurate instantaneous HR estimation with a very small estimation delay (3-4 s). The OSC-MSE-W algorithm, which allows the combination of multiple input signals and led to the smallest error values, was selected for further processing. We also showed that monitoring in the dark is possible with the investigated vision system. We noticed that, for the sequences recorded with an RGB camera, the robustness of HR estimation can be increased by combining the iPPG signals derived using different existing techniques. However, the analysis of the *motion* subset was challenging despite the simultaneous use of multiple iPPG signal derivation techniques supposedly resistant to motion [96, 105, 106]. An SQI relying on features derived from the images and the iPPG signal was developed in order to increase the accuracy of HR estimation when the subjects are moving. The resulting SQIs were used in the adaptive frequency tracking scheme, as well as to label and suppress very bad-quality segments, resulting in an error reduction of 42% for the *motion* subset.

To conclude, the proposed innovative approach takes advantage of different existing iPPG signal derivation techniques, by combining the relevant information present in each iPPG input signal in an adaptive frequency tracking scheme.

This processing scheme was further adapted to estimate HR from video sequences of preterm infants in the neonatal intensive care unit. Preliminary results can be found in Appendix B.

Finally, it is also worth mentioning that this video-based HR estimation scheme was implemented in C++ for a real-time demo⁸.

8. This demo was presented at Nano-Tera Final event 2017: <http://www.nano-tera.ch/events/meeting.html>

iPPG-Derived Pulse Rate Variability

9

9.1 Motivations

As mentioned in Section 7.1, iPPG is a very recent technology. There are therefore many aspects to investigate in order to have a better comprehension of this technology and its potential to replace traditional sensors (i.e. ECG-based sensors) in some cardiovascular monitoring applications.

The present chapter, which is divided into two main sections, focuses on iPPG-derived pulse rate variability (PRV). In the first part, agreement between ECG-HRV, PPG-PRV and iPPG-PRV indexes in non-stationary conditions induced by respiration and handgrip exercises is measured. In the second part, the iPPG-derived respiratory sinus arrhythmia is used to estimate the breathing frequency from the video sequences.

9.2 Assessment of iPPG pulse rate variability indexes

9.2.1 Introduction

The physiological phenomenon characterized by oscillations in the interval between consecutive heartbeats is known as heart rate variability (HRV). HRV has been extensively studied and some quantitative measures have been developed, which are now recognized as valuable measures of autonomic activity [75]. These measures are useful in clinical diagnostics, as they can be correlated with risk of cardiac events, such as congestive heart failure, sudden cardiac death and myocardial infarction [75, 133]. Usually, the R peaks detected on the ECG waveform are used to derive the inter-beat interval (IBI) time series, in this case called the RR intervals. In order to simplify the notation, the RR intervals will be noted ECG-IBI in the following analysis. The classical HRV indexes can be divided in two main classes: the time-domain measures and frequency-domain measures. Among the time-domain methods, the most used are the statistical methods, which consist of simple measures such as mean and standard deviation. Frequency domain measures focus on the analysis of the power present in different commonly used frequency bands: the ultra-low frequency (ULF) [0.0001-0.003] Hz, the very low frequency (VLF) [0.003-0.04] Hz, the low frequency (LF) [0.04-0.15] Hz and the high frequency (HF) [0.15-0.4] Hz bands. LF band is influenced both by the sympathetic and parasympathetic branches of the autonomic nervous system. The HF band is mainly influenced by the parasympathetic branch.

From a pulse-wave signal (PPG or ABP), one can derive the PRV. In that case, different pulse morphological landmarks can be used to compute the IBI time series, namely the systolic

peak, the maximum slope and the onset of the systole, as illustrated in Figure 9.1. Many studies have been performed to investigate whether the PRV indexes derived either from the PPG-IBI or ABP-IBI can be used as substitutes for HRV indexes computed from the ECG-IBI [134–146]. Different degrees of agreement have been reported in the aforementioned studies. Indeed, many factors have been found to influence the results. Among these factors, the most important are the experimental conditions (subject position, resting or non-resting condition), the pulse landmark used to compute the IBI series, and the sampling frequency. For example, Charlton *et al.* [136] found a good agreement between all the studied HRV parameters in supine position, while the agreement became poorer for the moderate and exhaustive exercise conditions. This decrease of agreement between PRV and HRV from resting to non-resting experimental conditions has also been observed in other studies [140, 143]. As pointed out by the authors of [134], the differences observed between the ECG-IBI and PPG-IBI time series are a consequence of two things: inaccuracies in pulse detection and physiological variability in the pulse transit time. The resulting PRV usually tends to overestimate the variability.

Nowadays, the same questions about the reliability of PRV indexes arise in the context of the iPPG technique. HR estimation using iPPG is usually investigated in the literature. However, the feasibility of performing PRV/HRV analysis using iPPG has not been investigated in many studies yet. In [100], one minute recordings were analyzed and the mean HR, as well as the LF, HF and LF/HF components were reported for iPPG signals derived from video sequences using ICA. The agreement was computed between PRV indexes derived from iPPG signals and PRV indexes derived from the reference finger blood volume pulse (BVP) sensor. An excellent agreement was reported for the average HR and a good agreement (correlation coefficient of about 0.9) for the frequency domain indexes. In [124], the relationship between HRV indexes derived from the ECG and PRV indexes derived from iPPG signals was studied. In this study, a database composed of 60 healthy subjects for which an autonomic response was triggered by a classical rest-to-stand manoeuvre was examined and the iPPG time series were derived from three different ROIs; one on the forehead, one on the nose and one on the cheek. The raw time series were first de-trended and the zero-phase component analysis (a variant of PCA) method was used to extract clean iPPG signals. The extraction of iPPG-IBI series was performed by finding peaks on the first derivative of the iPPG signal. The authors concluded that iPPG-PRV indexes can be considered as good surrogates for HRV indexes, especially in the supine position. The agreement between PRV and HRV parameters was lower in standing position, partly due to motion artifacts.

The aim of the present study is to assess the agreement between ECG-HRV, PPG-PRV and iPPG-PRV indexes in non-stationary conditions induced by respiration and handgrip exercises. For this purpose, a set of classical time domain and frequency domain HRV indexes were computed from the different IBI series and compared. As mentioned in Section 7.2, various methods have been proposed to derive the iPPG signals from the video frames. In this study, the iPPG signals derived using different techniques were compared, and the one leading to the most accurate IBI series was selected to compute PRV measures.

9.2.2 Data

The *respiration* and *handgrip* subsets of the iPPG database recorded on adult subjects described in Section 8.2 were used in this study.

9.2.3 Methods

The IBI series were computed from the three available waveforms: the ECG (ECG-IBI), the PPG (PPG-IBI) and the iPPG (iPPG-IBI). The accuracy of heartbeat detection on the PPG and iPPG signals was calculated and reported. The iPPG signal for which the heartbeat detection was



Figure 9.1: ECG and pulse-wave signal. The following pulse landmarks are indicated: the systolic peak (p), the maximum slope or first derivative (d) and the onset of the systole (f). An RR interval (RRI) is shown on the ECG, and the corresponding peak-to-peak interval (PPI(p)) is shown on the PPG.

the most accurate was selected for further analysis. Then, for each record, classical statistical and frequency domain HRV indexes were extracted from the three IBI series. Finally, the agreement between the HRV and PRV indexes was assessed using different measures. All methods are described into details in the following paragraphs.

Computation of ECG-IBI

The R-waves were detected on the ECG using a method for finding local maxima using a sliding window. The detected R-peaks were visually checked (and corrected if needed) in order to ensure the full correctness of our ground-truth.

Computation of PPG-IBI

The PPG baseline was roughly approximated by the zero-phase low-pass filtered version of the PPG (cutoff frequency of 0.8 Hz) and was subtracted from the original PPG. As mentioned earlier, different pulse landmarks can be used to compute the PPG-IBI time series. The main possibilities are the systolic peak, pulse onset and maximum slope location. Suhrbier *et al.* [135] showed that the maximum slope detection methods are more reliable than peak detection methods to compute the IBI series from a pulsatile signal, partly because the points detected using the former are more robust against morphological changes. For this reason, we decided to rely on an approach based on maximum slope detection for the computation of PPG-IBI and iPPG-IBI. More precisely, the technique used to detect ECG R-waves, based on local maxima detection, was used to detect maxima on the first PPG derivative (dPPG). Figure 9.2 illustrates the different steps of PPG heartbeats detection for a data segment taken from the *respiration* subset. Figure 9.3 shows an example of PPG signal corrupted by motion artifacts, resulting in heartbeat mis-detections.

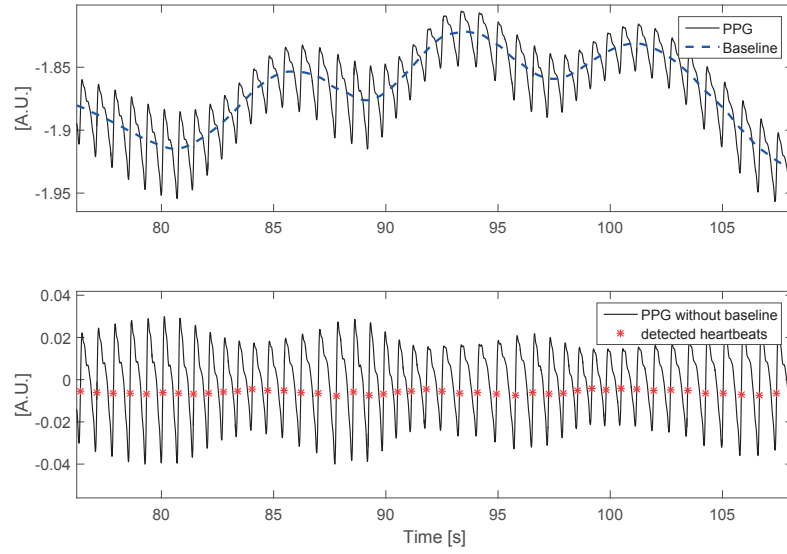


Figure 9.2: Exemple to illustrate PPG baseline removal and heartbeat detection.

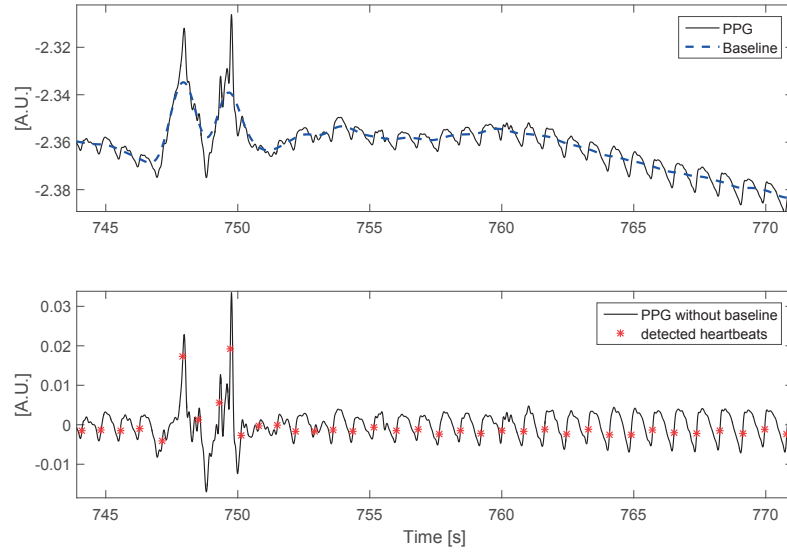


Figure 9.3: Example of a PPG segment corrupted by motion artifacts.

Computation of iPPG-IBI

The computation of iPPG-IBI series was the most challenging. Because of the contactless nature of this technology, the iPPG signals derived from the pixel intensities are prone to artifacts induced by subject movements and illumination changes. The morphology of the resulting waveforms is usually very variable compared to PPG. In order to have the same polarity as for the PPG signal, the sign of each iPPG signal was first compared to the sign of the PPG signal and,

if needed, the iPPG signal was inverted. Then, as for the PPG signal, heartbeats were detected using an approach based on maximum slope detection. In order to increase the detection accuracy, the derivative iPPG signal (diPPG) was first enhanced using the relative energy (Rel-En) algorithm described in [147, 148]¹. This algorithm extracts short- and long-term energies in a signal and provides a coefficient vector with which the signal is multiplied, heightening events of interest. More specifically, the ration between the short- and long-term energies is obtained trough:

$$c[n] = \frac{\sum_{i=n-s_{win}}^{n+s_{win}} |x[i]|^p}{\sum_{j=n-l_{win}}^{n+l_{win}} |w[j] \times x[j]|^p} \quad (9.1)$$

where the parameters s_{win} and l_{win} represent the half-lengths of the short and long sliding windows, respectively. The parameter p denotes the exponent and w is the window function of interest. The enhanced signal x_{RE} is computed as follows:

$$x_{RE}[n] = x[n]c[n] \quad (9.2)$$

An example of application presented in [147] is iPPG enhancement to improve heartbeat detection accuracy. The same parameters were used in the present study: 100 ms for the short window (s_{win}) and 600 ms duration for the long window (l_{win}). An exponent of 10 was selected for p . An example to illustrate the diPPG Rel-En enhancement and heartbeat detection is shown in Figure 9.4. A local cubic spline interpolation was finally performed for each detected maximum and its two closest neighbors using 200 points in order to increase the temporal resolution, as the sampling frequency of the iPPG signals was only 20 Hz. Heartbeat detection was performed on the different iPPG signals (iPPG_{green}, iPPG_{hue}, iPPG_{GRD}, iPPG_{SSR}, and iPPG_{POS}) that were derived using different techniques described in Section 7.2.

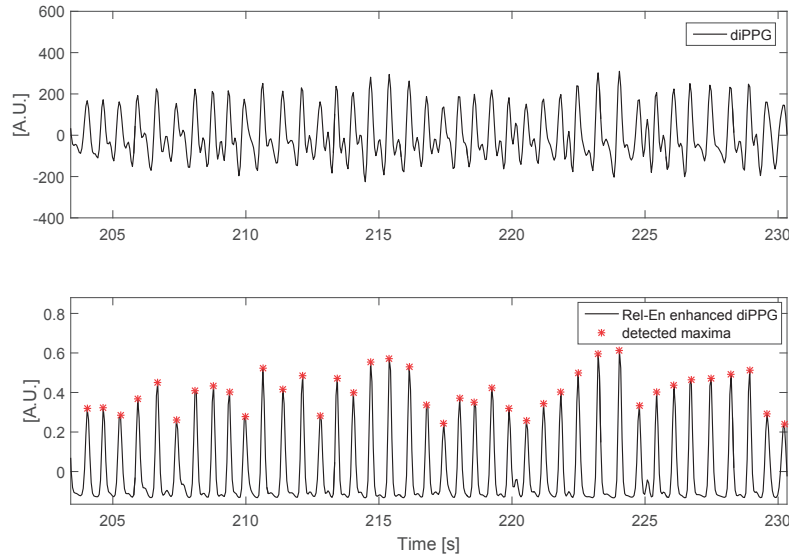


Figure 9.4: Example to illustrate heartbeat detection on the iPPG signal. The original diPPG signal (on top) is shown, as well as the RelEn enhanced diPPG signal (bottom), on which the maxima were detected.

1. This algorithm was developed by Sasan Yazdani (ASPG, EPFL).

Figures 9.5 and 9.6 show examples of the computed IBI series for sequences 1 and 6, respectively.

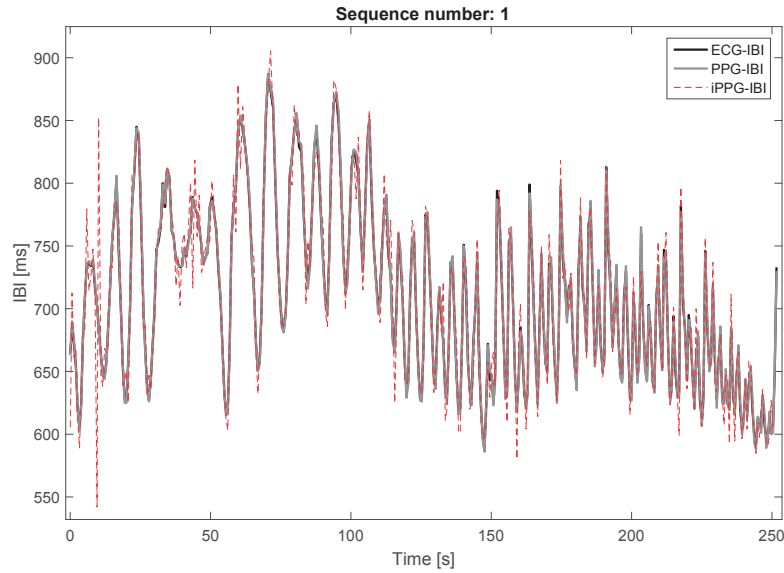


Figure 9.5: The three IBI series for a subject (sequence 1, *respiration* protocol). The iPPG-IBI series showed in this example was derived from the iPPG_{SSR} signal.

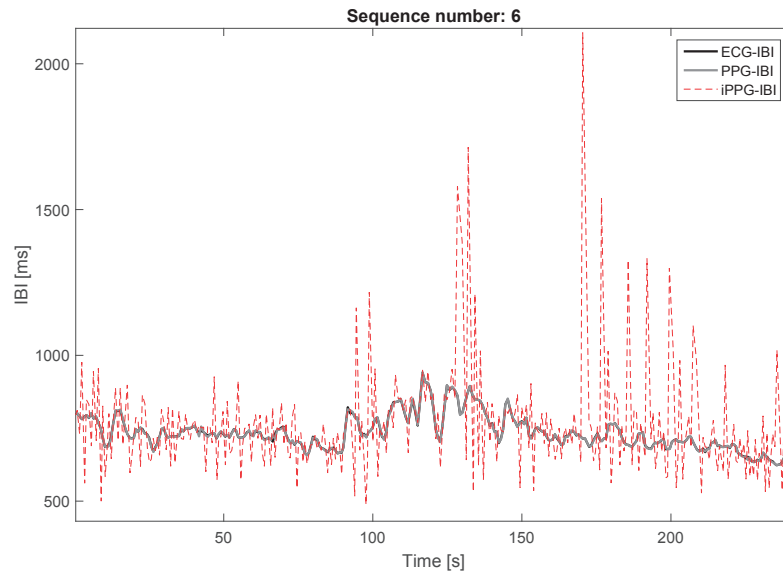


Figure 9.6: The three IBI series for a subject (sequence 6, *handgrip* protocol). The iPPG-IBI series showed in this example was derived from the iPPG_{SSR} signal.

Accuracy of heartbeat detection

The detected heartbeats from the PPG and iPPG waveforms were first aligned with the detected heartbeats on ECG by compensating for the average pulse transit time. For each sequence, the latter was computed as the mean distance between an R-wave and the closest subsequent PPG/iPPG detected heartbeat. In order to assess the heartbeat detection accuracy, the heartbeats detected in the PPG/iPPG signals were classified into three categories: true positive (TP), false positive (FP) and false negative (FN). A tolerance of 150 ms was used. Then, the following performance metrics were calculated:

The detection error rate (DER):

$$DER = \frac{FP + FN}{N_{peaksref}} \quad (9.3)$$

The sensitivity (Se):

$$Se = \frac{TP}{TP + FN} \quad (9.4)$$

The positive prediction value (PPV):

$$PPV = \frac{TP}{TP + FP} \quad (9.5)$$

where $N_{peaksref}$ indicates the total number of heartbeats in the reference signal, i.e. the ECG.

Computation of the HRV/PRV indexes

Some of the classical HRV indexes were computed [75]. In the time domain:

- Mean HR [bpm]: 60 divided by the mean of all normal-to-normal (NN) intervals².
- SDNN [ms]: Standard deviation of all NN intervals.
- RMSSD [ms]: The square root of the mean of the sum of the squares of differences between adjacent NN intervals.
- pNN50 [%]: NN50 count divided by the total number of all NN intervals. NN50 is the number of interval differences of successive NN intervals greater than 50 ms.

In the frequency domain: The IBI series were first uniformly re-sampled at 4 Hz. A non-parametric method (Welch periodogram) was used to compute the PSD.

- LF norm [%]: Power in the LF band in normalized units. Frequency range [0.04-0.15] Hz.
- HF norm [%]: Power in the HF band in normalized units. Frequency range [0.15-0.4] Hz.
- LF/HF : Ratio LF [ms²] /HF [ms²].

Agreement between HRV and PRV parameters:

The Pearson correlation coefficient (r) between the HRV and PRV parameters was computed, as this coefficient is usually reported in previous studies. However, it should be noted that this coefficient can only assess the linear correlation between the studied parameters, but yields an incomplete information. For example, the presence of a systematic bias can lead to a large correlation coefficient. For this reason, a Bland-Altman analysis might be more appropriate. More specifically, the Bland-Altman ratio (BA_r) can be computed as:

$$BA_r = 0.5 \frac{CI_{95\%}}{mean(AV_{(ref,est)})} \quad (9.6)$$

2. NN intervals: Intervals between adjacent QRS complexes resulting from sinus node depolarization.

with $CI_{95\%}$ the 95% confidence interval of the differences between the reference and the estimation and $AV_{(ref,est)}$, all the average values of the two measures. The agreement can be qualified as excellent when $BA_r < 0.01$, good if $BA_r \in [0.01 - 0.1]$, moderate if $BA_r \in [0.1 - 0.2]$ and insufficient if $BA_r > 0.2$.

For each HRV/PRV index, and for each sequence, the difference between reference and the measure of interest (referred as Δ) was also computed and tailed Wilcoxon signed rank tests were performed to test the alternative hypothesis stating that the data Δ come from a distribution with median greater/smaller than 0. The significance level was set at 1%.

9.2.4 Results

9.2.5 Heartbeat detection accuracy

Regarding the heartbeat detection accuracy from the PPG waveforms, the average DER was very small (0.008 ± 0.018), and was zero for most of the records (Table 9.1). A few mis-detections occurred in some records, caused by motion artifacts. Figure 9.3 shows an example of PPG signal corrupted by motion artifacts.

Table 9.1: Performance of heartbeat detection on the PPG signal, for each sequence.

Record number	Protocol	DER	Se	PPV
1	<i>respiration</i>	0	1	1
2	<i>handgrip</i>	0	1	1
3	<i>respiration</i>	0	1	1
4	<i>handgrip</i>	0.004	0.996	1
5	<i>respiration</i>	0	1	1
6	<i>handgrip</i>	0	1	1
7	<i>respiration</i>	0	1	1
8	<i>handgrip</i>	0	1	1
9	<i>respiration</i>	0	1	1
10	<i>respiration</i>	0	1	1
11	<i>handgrip</i>	0	1	1
12	<i>respiration</i>	0	1	1
13	<i>handgrip</i>	0.026	0.987	0.987
14	<i>respiration</i>	0.073	0.963	0.963
15	<i>handgrip</i>	0.015	0.994	0.991
16	<i>respiration</i>	0.003	1	0.997
17	<i>handgrip</i>	0	1	1
18	<i>respiration</i>	0	1	1
19	<i>respiration</i>	0.004	1	0.996
20	<i>handgrip</i>	0.044	0.972	0.984
21	<i>respiration</i>	0	1	1
22	<i>handgrip</i>	0	1	1
Average <i>respiration</i> sequences		0.007 ± 0.021	0.997 ± 0.011	0.996 ± 0.011
Average <i>handgrip</i> sequences		0.009 ± 0.015	0.995 ± 0.009	0.996 ± 0.006
Average all sequences		0.008 ± 0.018	0.996 ± 0.010	0.996 ± 0.009

The mean DER, Se and PPV for the different iPPG signals are reported in Table 9.2. It can be noted that, on average, the iPPG_{SSR} signal led to smallest DER. The iPPG_{SSR} signal was therefore selected to perform further analysis.

Table 9.2: Mean heartbeat detection performance metrics for all sequences and the different iPPG signals.

Input iPPG	average DER	average Se	average PPV
iPPG _{green}	0.095 ± 0.079	0.946 ± 0.049	0.958 ± 0.035
iPPG _{hue}	0.131 ± 0.168	0.928 ± 0.098	0.939 ± 0.080
iPPG _{GRD}	0.083 ± 0.079	0.955 ± 0.049	0.962 ± 0.036
iPPG _{SSR}	0.078 ± 0.086	0.956 ± 0.053	0.965 ± 0.038
iPPG _{POS}	0.110 ± 0.089	0.943 ± 0.051	0.947 ± 0.043

Table 9.3 shows the heartbeat detection performance metrics on the iPPG_{SSR} signal, for each sequence.

Table 9.3: Performance of heartbeat detection on the iPPG_{SSR} signal, for each sequence.

Record number	Protocol	DER	Se	PPV
1	<i>respiration</i>	0.011	0.992	0.997
2	<i>handgrip</i>	0.084	0.937	0.978
3	<i>respiration</i>	0.033	0.991	0.977
4	<i>handgrip</i>	0.070	0.971	0.960
5	<i>respiration</i>	0.087	0.952	0.961
6	<i>handgrip</i>	0.199	0.882	0.915
7	<i>respiration</i>	0.013	0.990	0.997
8	<i>handgrip</i>	0.011	0.993	0.996
9	<i>respiration</i>	0.063	0.965	0.972
10	<i>respiration</i>	0.143	0.925	0.932
11	<i>handgrip</i>	0.185	0.906	0.909
12	<i>respiration</i>	0.022	0.984	0.994
13	<i>handgrip</i>	0.051	0.968	0.980
14	<i>respiration</i>	0.006	0.994	1.000
15	<i>handgrip</i>	0.378	0.762	0.845
16	<i>respiration</i>	0.058	0.966	0.976
17	<i>handgrip</i>	0.014	0.990	0.997
18	<i>respiration</i>	0.041	0.978	0.981
19	<i>respiration</i>	0.100	0.946	0.953
20	<i>handgrip</i>	0.056	0.972	0.972
21	<i>respiration</i>	0.072	0.988	0.943
22	<i>handgrip</i>	0.019	0.989	0.992
Average <i>respiration</i> sequences		0.054 ± 0.042	0.973 ± 0.022	0.973 ± 0.022
Average <i>handgrip</i> sequences		0.107 ± 0.116	0.937 ± 0.072	0.954 ± 0.050
Average all sequences		0.078 ± 0.086	0.956 ± 0.053	0.965 ± 0.038

Tables 9.4 and 9.5 report the computed measures of agreement between HRV and PPG-PRV indexes for the *respiration* and *handgrip* sequences, respectively.

Table 9.4: Summary of the agreement between the HRV and PPG-PRV indexes for *respiration* sequences. * indicates that the p-value was lower than 0.01.

	r	BAr	av. $\Delta \pm \text{std}$, (tested hypothesis)
av. HR	1.000	0.008	-0.158 ± 0.291 [bpm], ($\Delta > 0$)
SDNN	0.940	0.236	$-4.792 \pm 9.762^*$ [ms], ($\Delta < 0$)
RMSSD	0.782	0.699	$-9.950 \pm 21.901^*$ [ms], ($\Delta < 0$)
pNN50	0.945	0.445	$-3.717 \pm 6.601^*$ [%], ($\Delta < 0$)
LF	0.650	0.172	$2.200 \pm 4.073^*$ [%], ($\Delta > 0$)
HF	0.675	0.350	$-3.483 \pm 6.441^*$ [%], ($\Delta < 0$)
LF/HF	0.837	0.402	$0.162 \pm 0.283^*$, ($\Delta > 0$)

Table 9.5: Summary of the agreement between the HRV and PPG-PRV indexes for *handgrip* sequences. * indicates that the p-value was lower than 0.01.

	r	BAr	av. $\Delta \pm \text{std}$, (tested hypothesis)
av. HR	1.000	0.008	-0.030 ± 0.295 [bpm], ($\Delta < 0$)
SDNN	0.946	0.295	$-7.800 \pm 13.398^*$ [ms], ($\Delta < 0$)
RMSSD	0.569	1.304	$-22.590 \pm 37.207^*$ [ms], ($\Delta < 0$)
pNN50	0.949	0.410	-2.420 ± 3.874 [%], ($\Delta < 0$)
LF	0.525	0.214	$2.900 \pm 5.561^*$ [%], ($\Delta > 0$)
HF	0.592	0.889	$-5.430 \pm 9.559^*$ [%], ($\Delta < 0$)
LF/HF	0.792	0.738	$0.835 \pm 1.142^*$, ($\Delta > 0$)

Tables 9.6 and 9.7 report the computed measures of agreement between HRV and iPPG-PRV indexes for the *respiration* and *handgrip* sequences, respectively.

Table 9.6: Summary of the agreement between the HRV and iPPG-PRV indexes for *respiration* sequences. * indicates that the p-value was lower than 0.01.

	r	BAr	av. $\Delta \pm \text{std}$, (tested hypothesis)
av. HR	0.995	0.033	$-0.783 \pm 1.200^*$ [bpm], ($\Delta < 0$)
SDNN	0.807	0.416	$-24.667 \pm 19.345^*$ [ms], ($\Delta < 0$)
RMSSD	0.675	0.806	$-65.133 \pm 36.610^*$ [ms], ($\Delta < 0$)
pNN50	0.644	0.751	$-26.242 \pm 15.473^*$ [%], ($\Delta < 0$)
LF	0.524	0.243	$9.375 \pm 5.292^*$ [%], ($\Delta > 0$)
HF	0.676	0.363	$-13.550 \pm 7.619^*$ [%], ($\Delta < 0$)
LF/HF	0.570	0.668	$0.595 \pm 0.397^*$, ($\Delta > 0$)

Table 9.7: Summary of the agreement between the HRV and iPPG-PRV indexes for *handgrip* sequences. * indicates that the p-value was lower than 0.01.

	r	BAr	av. $\Delta \pm \text{std}$, (tested hypothesis)
av. HR	0.996	0.044	0.290 ± 1.646 [bpm], ($\Delta > 0$)
SDNN	0.234	0.980	$-45.150 \pm 53.827^*$ [ms], ($\Delta < 0$)
RMSSD	-0.408	1.861	$-106.640 \pm 93.003^*$ [ms], ($\Delta < 0$)
pNN50	0.210	1.023	$-38.950 \pm 19.187^*$ [%], ($\Delta < 0$)
LF	0.323	0.416	$13.280 \pm 9.722^*$ [%], ($\Delta > 0$)
HF	0.329	1.053	$-24.230 \pm 16.369^*$ [%], ($\Delta < 0$)
LF/HF	0.328	1.498	$2.293 \pm 1.760^*$, ($\Delta > 0$)

Tables 9.8 and 9.9 report the computed measures of agreement between PPG-HRV and iPPG-PRV indexes for the *respiration* and *handgrip* sequences, respectively.

Table 9.8: Summary of the agreement between the PPG-PRV and iPPG-PRV indexes for *respiration* sequences. * indicates that the p-value was lower than 0.01.

	r	BAr	av. $\Delta \pm \text{std}$, (tested hypothesis)
av. HR	0.994	0.036	-0.625 ± 1.303 [bpm], ($\Delta < 0$)
SDNN	0.746	0.457	$-19.875 \pm 21.808^*$ [ms], ($\Delta < 0$)
RMSSD	0.440	0.966	$-55.183 \pm 46.306^*$ [ms], ($\Delta < 0$)
pNN50	0.541	0.774	$-22.525 \pm 16.684^*$ [%], ($\Delta < 0$)
LF	0.168	0.352	$7.175 \pm 7.487^*$ [%], ($\Delta > 0$)
HF	0.296	0.518	-10.067 ± 11.337 [%], ($\Delta < 0$)
LF/HF	0.322	0.952	0.433 ± 0.526 , ($\Delta > 0$)

Table 9.9: Summary of the agreement between the PPG-PRV and iPPG-PRV indexes for *handgrip* sequences. * indicates that the p-value was lower than 0.01.

	r	BAr	av. $\Delta \pm \text{std}$, (tested hypothesis)
av. HR	0.995	0.048	0.320 ± 1.823 [bpm], ($\Delta > 0$)
SDNN	0.189	0.984	-37.350 ± 56.039 [ms], ($\Delta < 0$)
RMSSD	-0.181	1.807	$-84.050 \pm 100.721^*$ [ms], ($\Delta < 0$)
pNN50	0.218	1.007	$-36.530 \pm 19.507^*$ [%], ($\Delta < 0$)
LF	0.328	0.451	$10.380 \pm 10.203^*$ [%], ($\Delta > 0$)
HF	0.348	1.017	$-18.800 \pm 17.216^*$ [%], ($\Delta < 0$)
LF/HF	0.358	1.346	$1.458 \pm 1.295^*$, ($\Delta > 0$)

Figures 9.7 and 9.8 show the scatter plots for all the computed HRV parameters, as well as the Pearson correlation coefficient (r) and the BAr , for PPG and iPPG, respectively.

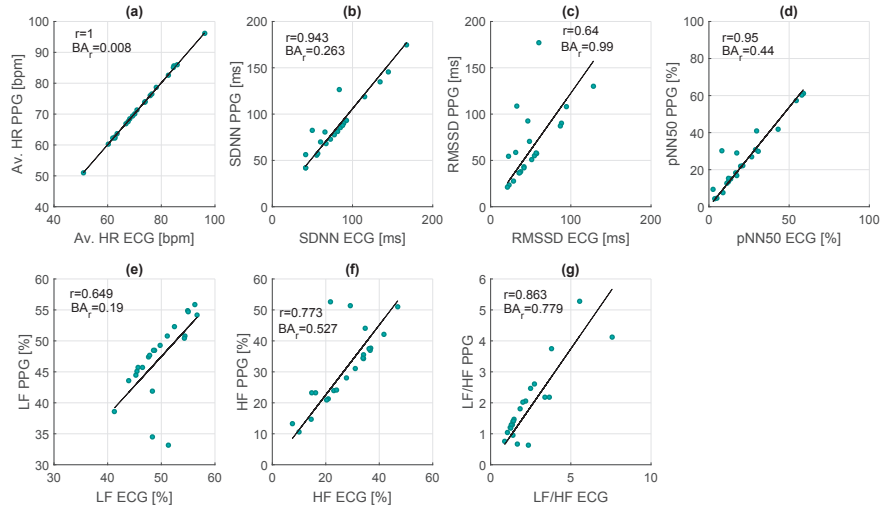


Figure 9.7: Agreement between ECG-HRV and PPG-PRV time domain measures [(a)-(d)] and frequency domain measures [(e)-(g)], for all sequences.

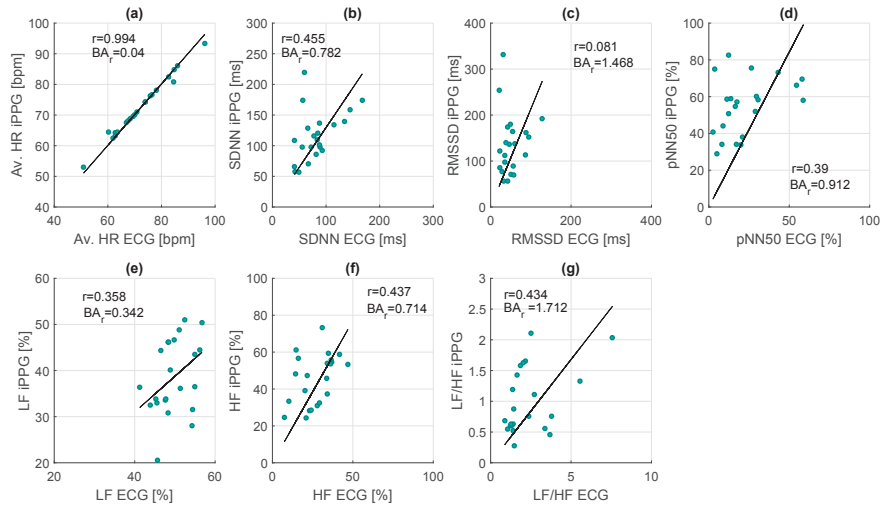


Figure 9.8: Agreement between ECG-HRV and iPPG-PRV time domain measures [(a)-(d)] and frequency domain measures [(e)-(g)], for all sequences.

9.2.6 Discussion

With an average DER of 0.008 ± 0.018 , Se of 0.996 ± 0.010 and PPV of 0.996 ± 0.009 , the accuracy of heartbeat detection from the PPG signal was good. Some mis-detections occurred, mainly due to the motion artifacts induced by the physical effort required to perform the handgrip exercise (Figure 9.3). It can be seen in Tables 9.4 and 9.5 that the best agreement was obtained for the average HR. This makes sense, because the little variations tend to cancel each other in the averaging process. The pNN50 and SDNN were the second best indexes in terms of agreement, for both the *respiration* and *handgrip* subsets. Despite the very low DER obtained for PPG heartbeat detection, it can be noticed that the level of agreement was not very high for some indexes. More precisely, r was lower than 0.9 for the RMSSD and all the frequency domain indexes and BA_r was higher than 0.2 for most of the studied parameters. The Δ measure on the time domain statistical indexes indicates that the variability was overestimated. Indeed, the $SDNN_{PPG}$, $RMSSD_{PPG}$, $PNN50_{PPG}$ were significantly higher than $SDNN_{ECG}$, $RMSSD_{ECG}$, $PNN50_{ECG}$. These trends are in accordance with previous findings [134]. Regarding the difference between the type of exercise performed by the subjects, it can be noted that, on average and for all parameters, a higher degree of agreement was obtained for the *respiration* subset. This might be partly explained by the observed difference in heartbeat detection accuracy. However, it is important to emphasize that the observed bias between HRV and PRV is not entirely explained by fluctuations in peak detection but is also related to physiological effects. Indeed, the pulse transit time (PTT) is known to have some variability. In [149], the PTT changes induced by paced respiration activity were studied and the authors concluded that a strong relationship between PTT changes and RR intervals changes is present, but changes are not synchronous.

Regarding the iPPG signals, the best results (DER of 0.078 ± 0.086 , Se of 0.956 ± 0.053 and PPV of 0.965 ± 0.038) were obtained for the iPPG_{SSR} signal. The same performance metrics yielded slightly better results for the *respiration* subset, which can be again partially explained by the presence of motion artifacts in the *handgrip* subset. Tables 9.6 and 9.7 show the agreement between ECG-HRV and iPPG-PRV. As for PPG, the highest agreement was obtained for the mean HR. For all the studied indexes, the agreement was lower than that of PPG-derived indexes, which is not really surprising given the challenging nature of the iPPG technology. As for PPG, a higher degree of agreement was obtained for the *respiration* subset. The obtained r values for the respiration subset are very close to the ones obtained on iPPG signals in [124] for a standing dataset, for which the performance of heartbeat detection were also very similar to ours. The level of agreement between iPPG-PRV and ECG-HRV indexes was very low for the *handgrip* subset, for all the studied indexes except the average HR. For both PPG and iPPG, it can be seen that the relative HF power was significantly increased compared to the ECG-derived one. These results are in accordance with previous findings [144, 150] and suggest that, because of this overall low level of agreement for frequency domain indexes, both PPG-PRV and iPPG-PRV are not suitable to study autonomic cardiovascular control. Tables 9.8 and 9.9 report the agreement between PPG-PRV and iPPG-PRV, which is surprisingly lower than the agreement between ECG-HRV and iPPG-PRV, for all variability measures and for the two data subsets. This suggests that the error/variability sources were quite different.

Different conclusions were drawn in previous studies regarding the information contained in the PRV derived either from PPG or ABP waveforms. The results, which are not unequivocal, are nicely summarized by Schäfer and Vagedes [134]. Among various opinions, most of the authors warned against the use of PRV as a surrogate of HRV in specific conditions. For example, Charlot *et al.* [136] did not recommend the use of PRV, except for the supine rest setting. The obtained results point in the same direction, and show that one should be very careful when using iPPG to study HRV, as the level of agreement is even lower than that of the PPG-derived PRV indexes.

It was shown in the present study that iPPG-PRV is neither a good surrogate of ECG-HRV nor of PPG-PRV in non-resting conditions. However, it is important to emphasize that the method used to derive the iPPG signal plays an important role. The results presented in this chapter were generated using the actual state-of-the-art iPPG signal derivation techniques. These techniques might considerably evolve in the near future and become more robust and subsequently lead to results closer to the ones obtained with the PPG signal. Another factor which might have lowered the agreement of iPPG-derived indexes compared to PPG is the low sampling frequency of 20 Hz, despite the interpolation performed. In addition, it would be interesting to see if trends are conserved. For example, a patient goes from a state A to a state B, which is characterized by a significantly higher variability, will the variability derived from iPPG be significantly higher too? If yes, it makes iPPG-PRV of great value. For instance, the detection of atrial fibrillation (AF) using iPPG was investigated by Couderc *et al.* [151]. In this study, a novel measure of PRV called the pulse harmonic strength has been shown to be able to differentiate between different states (i.e. sinus rhythm and AF).

To conclude, it would be interesting to perform the same analysis on a similar database with subjects at rest in order to see if the level of agreement between ECG-HRV and iPPG-PRV increases. Unfortunately, such a database was not available in the context of this work. However, another aspect of iPPG PRV was briefly studied, which is estimation of the respiratory rate by using the iPPG PRV.

9.3 Real-time respiratory rate estimation using imaging photoplethysmography inter-beat intervals

9.3.1 Introduction

iPPG waveforms can also be used to estimate respiratory information directly [125, 152] or indirectly through the respiratory modulation of the IBI. The latter is usually measured as the HF component of the HRV [100]. These studies employ spectral estimation or empirical mode decomposition, which require analyzing fixed bandwidths and/or fixed-length segments of the iPPG waveforms. Fixed bandwidths, namely the HF 0.15-0.4 Hz band [153] are restrictive and result in erroneous estimations when respiratory rates are lower than 0.15 Hz (9 breaths-per-minute) and fixed-length segments analysis is detrimental when considering real-time applications. In the present study, we aim to estimate the respiratory rate from the iPPG signal (acquired in a contactless manner) by using the PRV, and a real-time-capable algorithm, without restrictive bandwidths. We used the instantaneous and real-time NFB algorithm (briefly described in Section 8.3.3) [131, 154] to track the main frequency component of the PRV, which is the respiration. We compared our estimates to an estimate computed from the simultaneously recorded ECG, which is an accepted method of estimating the respiratory rate from cardiac activity [155, 156]. This study was the object of a conference publication [157]³.

9.3.2 Methods

In this study, the iPPG_{red}, iPPG_{green} and iPPG_{blue} waveforms were obtained by the averaging signal derivation method described in Section 7.2, using a fixed ROI extracted from the forehead of the subjects. The *respiration* and *handgrip* subsets were described in Section 8.2 were used. The instantaneous heart rate was computed by using the NFB algorithm [131, 154] with the three channels as simultaneous inputs. The mean heart rate was used to compute an appropriate inter-beat estimate for an extrema detection method to identify the heart beats in each channel. The

3. This study was done in collaboration with Leila Mirmohamadsadeghi, former PhD student of EPFL-ASPG.

IBI time series were created for each channel and all three series were re-sampled uniformly at a sampling rate of 4 Hz using linear interpolation to avoid delays [158]. The re-sampled series were then band-pass filtered between 0.09 and 1 Hz, considered to be a large respiratory band comprising rates from 5.4 to 60 breaths-per-minute. The respiratory rate was estimated using a modified version of the NFB algorithm with each of the three filtered iPPG IBI as input. This modified version (compared to [131, 154]) comprises an extra pre-processing step in which the sliding window singular value decomposition (SWASVD) algorithm [128] (see Section 8.3.3) was applied to extract the principal component of the IBI series. This step was added as the wide-band filtering does not produce an oscillatory signal and thus might result in erroneous estimation through the classic NFB. For comparison, the respiratory rate was similarly estimated from the ECG IBI.

The reference respiratory rate was computed from the reference respiratory signal acquired with the impedance belt. This signal was re-sampled uniformly at 4 Hz and filtered similarly to the iPPG and ECG IBI. Its instantaneous frequency was then estimated in three ways: (1) by identifying the largest peak of the Welch spectrum in sliding windows, (2) using the NFB, and (3) the average of (1) and (2).

The correlations between the iPPG and ECG IBI were computed with Pearson's correlation coefficient, r . The errors of their smoothed (4 s windows) estimates were computed as the mean absolute difference, in breaths-per-minute, between the estimates and the smoothed (4 s windows) reference rate. However, due to large artifacts in the iPPG, it was necessary to develop a quality index to identify portions of sufficient quality and to consider errors only in good quality segments. Indeed, in practice, one would prefer to know that the signals are of poor quality rather than being presented with a bad estimate. Therefore, an empirical quality index based on the amplitude of the iPPG signals was developed. To compute this quality index, an amplitude index was computed as the squared amplitude of the signal, divided by its variance (computed in a sliding window). The quality index was set to 1 where the amplitude index was smaller than ten times its interquartile range (computed in a sliding window) and 0 elsewhere. Portions with a quality index of 1 were retained when computing the correlations and the errors.

9.3.3 Results

The evaluation was performed on the entire length of the recordings over all parts of the protocols where the quality index was equal to 1. Among the three iPPG channels, the green channel yielded the IBI most similar to the ECG ones with an average correlation of 0.65 ± 0.27 (all per-record correlations were significant with $p < 0.001$). The red and blue channels intervals yielded correlations with the ECG intervals below 0.5 as reported in Table 9.10. Therefore the green channel was retained for respiratory rate estimation.

Table 9.10: The mean correlations between the IBI of the three iPPG channels and the ECG inter-beat intervals over all subjects.

	iPPG _{red}	iPPG _{green}	iPPG _{blue}
ECG	0.04 ± 0.11	0.65 ± 0.27	0.38 ± 0.24

Figures 9.9 and 9.10 illustrate the iPPG and ECG respiratory rate estimates and the reference for two subjects. The iPPG quality index is shown as well. It can be seen that both estimates follow the steep increase in the reference for both subjects.

The errors between the estimates and the reference are reported in Table 9.11 for both the ECG and the iPPG. The difference between the ECG and iPPG estimates was 3.05 ± 1.69 . The results were obtained on 88% of the data selected as having sufficient iPPG quality.

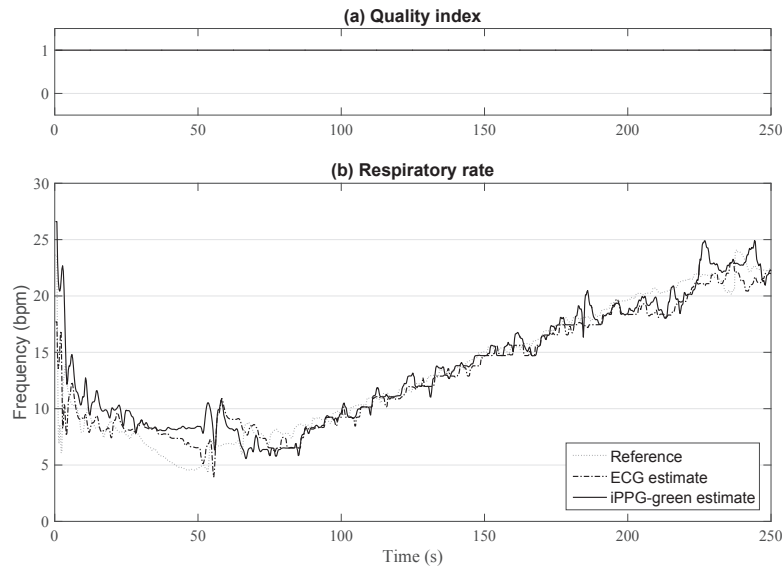


Figure 9.9: The iPPG_{green} (a) quality index and (b) respiratory rate, ECG-derived respiratory rate and reference for subject 1.

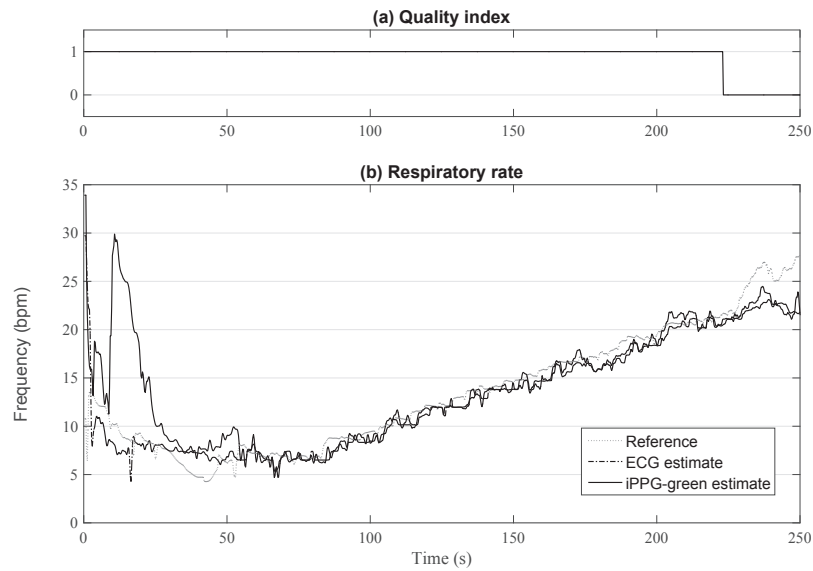


Figure 9.10: The iPPG_{green} (a) quality index and (b) respiratory rate, ECG-derived respiratory rate and reference for subject 2.

Table 9.11: The mean (\pm standard deviation) error in breaths-per-minute of the ECG and iPPG_{green} estimates compared to the reference over all subjects.

	Welch ref. (1)	NFB ref. (2)	mean ref. (3)
ECG	3.18 ± 2.94	2.86 ± 3.67	2.74 ± 3.25
iPPG _{green}	4.06 ± 1.88	3.49 ± 2.32	3.52 ± 1.99

9.3.4 Discussion

The iPPG_{green} signal yielded the IBI most similar to the ECG ones, which corroborates earlier findings on the green wavelength being more suitable than red and blue to capture skin tone differences due to blood circulation because of its better absorption by hemoglobin [5]. The accuracy of the iPPG estimates was slightly less than that of ECG estimates. However, considering the challenging conditions of the iPPG acquisition and processing, the errors are still comparable to errors reported in the literature for estimating the respiratory rate from the ECG or the PPG [155, 156].

Regarding the apnea detection, it has been shown that the respiratory influence on the IBI subsists in a weaker form even without respiration [159], which means that apneas cannot be detected reliably via the IBI. Therefore the subject of apnea detection with iPPG-derived IBI was not addressed in the present study.

Limitations: In the present study, the subjects were in a supine position. In this position, most often the respiratory influence on the HRV is much stronger than the baroreflex activity, occurring at 0.1 Hz. Therefore, the main component of the HRV in all but one subject was the respiration. However, in an orthostatic position, the baroreflex activity would be larger than that of the respiration and the previous assumption would not hold. It should also be mentioned that, despite our best care and intentions, the reference respiratory rates were prone to errors and artifacts and do not represent a ground-truth.

9.4 Conclusion

The iPPG-derived PRV was studied in this chapter. In the first part, classical HRV measures were derived from the ECG, PPG and iPPG waveforms and their level of agreement was evaluated. It was shown that iPPG-PRV is neither a good surrogate of ECG-HRV nor of PPG-PRV in non-resting conditions. However, even if PPG-PRV and iPPG-PRV were different from the HRV computed from the ECG, if the same trend of differences is observed between various physiological conditions, it makes the PPG-PRV and iPPG-PRV of great value. It will therefore be important to perform further investigations in that regard. In the second part, the iPPG PRV was used to estimate the respiratory rate. The obtained estimation errors were comparable to commonly reported errors for respiratory rate estimation from the ECG and the conventional contact-based PPG. Moreover, the data of the *respiration* subset were recorded with varying respiration rates, which is challenging. These findings are encouraging in the use of iPPG for real-time contactless respiration rate monitoring.

Conclusion

The main focus of this thesis was the development of processing tools for cardiovascular monitoring applications based on conventional and video-based PPG. Because of their ability to monitor comfortably the subjects during their daily life activities, PPG-based wearable monitoring devices are becoming very popular. Besides, imaging PPG has also attracted much attention in the last years. However, the PPG signals acquired by wearable devices, or remotely using a video-camera, are easily deteriorated by subject movements, illumination changes (iPPG) or by a weak contact between the probe and the tissue (PPG). Thus, efforts are made by the engineers to 1) develop new processing tools to improve the reliability of the estimated parameters and 2) explore the different possible applications of this technology. These two research areas were studied in this thesis.

One part of this work focused on improving the reliability of the heart rate (HR) estimated using a wrist-type PPG device. After introducing the principle of PPG, we presented the challenges related to the presence of motion artifacts on PPG signals and the state-of-the-art methods to deal with these motion artifacts. We then introduced a new straightforward method to estimate HR from subjects performing different kind physical activities. In the proposed scheme, the accelerometer signals were used to performed adaptive noise attenuation on the PPG signals, using a normalized least-mean-square algorithm. An adaptive band-pass filter was then used to track the instantaneous frequency (i.e. HR) of the reconstructed PPG signals. These adaptive frequency tracking tools were previously developed at the ASPG by Jérôme Van Zaen [23] and Yann Prudat [22], former PhD researchers of the ASPG. We developed an additional mechanism to have a time-varying parameters controlling the convergence speed and the band-width of the adaptive band-pass filter. This mechanism was based on the spectral purity index. We reported the performance of the proposed approach on the publicly available *IEEE Signal Processing Cup 2015* database. The results obtained were in the same range than the results reported in previous studies on the same database. However, unlike the previously proposed Fourier-based approaches, our scheme provided continuous HR estimates with a very low time delay, which is of great interest for real-time applications. In addition, all the signals were automatically combined in an optimal way and thus, no a priori knowledge about the contribution of the different waveforms was required. Finally, only a very few number of parameters had to be tuned, making this approach suitable and flexible for different experimental situations.

Another part of this work on PPG signals consisted of studying the feasibility of detecting atrial fibrillation (AF) from PPG signals acquired with a wrist-type device. This study took place in the context of the Nano-Tera NTF project MiniHolter, in collaboration with the Swiss Center

for Electronics and Microtechnology (CSEM) and the Lausanne University Hospital (CHUV). PPG signals recorded during cardiac ablation procedures were analyzed. Based on the simple observation that the PPG waveform is a lot more irregular in terms of morphology and amplitude during AF episodes, we investigated features able to reflect such changes. These features, namely the adaptive organization index, the variance of the slope of the phase difference between fundamental and first harmonic components, the permutation entropy, the spectral entropy, the fractional spectral radius and the spectral purity index, are innovative in the context of PPG processing. The adaptive organization index and the variance of the slope of the phase difference features were previously proposed by a former PhD researcher of ASPG (Andréa Buttu) for the analysis of surface ECG atrial activity during AF [160]. Other features derived from the PPG interbeat intervals (IBI) were also computed. The annotated ECG signals were used as ground-truth to assess the performance of the proposed features. We investigated not only the discrimination between AF and sinus rhythm (SR), but also between AF and ventricular arrhythmias (VA). We compared these PPG-wave features with the PPG-IBI features and concluded that they had a similar discriminative capability to discriminate between SR and AF, and a higher discriminative capability to discriminate between AF and VA. We also pointed the fact the high observed variability in PPG-IBI features during AF episodes was not only a result of AF but also a consequence of a poorer heartbeat detection accuracy. We concluded that the proposed PPG-wave features have a good potential to be used as complementary measures to the more classical PPG-IBI features.

The reduction of false arrhythmia alarms using signals from independent sources, namely PPG, ECG and arterial blood pressure (ABP), was also studied in this thesis. This study took place in the context of *PhysioNet/CinC challenge 2015*. The *PhysioNet/CinC challenge 2015* database consists of 1200 life-threatening alarms recordings, each categorized as a bradycardia, tachycardia, asystole, ventricular tachycardia or ventricular flutter/fibrillation. In the proposed approach, the quality of the pulsatile waveforms (PPG and ABP) was first assessed using a publicly available algorithm. Then, based on the quality of available signals, HR was either estimated from pulsatile waveforms using an adaptive frequency tracking algorithm or computed from ECGs using an adaptive mathematical morphology approach. We also introduced a measure based on the spectral purity of the ECGs to determine if a ventricular tachycardia or flutter/fibrillation has taken place. Finally, alarm veracity was determined based on a set of decision rules on HR and spectral purity values. The proposed approach, developed in collaboration with Sasan Yazdani (PhD researcher of the ASPG), was tested on a hidden dataset by the organizers of this challenge. This resulted in overall true positive rates of 95%/99% and overall true negative rates of 76%/80% on the real-time and retrospective subsets of the test dataset, respectively. In addition to these good results (the first prize for the “retrospective” dataset), we believe that the proposed approach is interpretable for the medical doctors, as we decided to keep the decision making process straightforward by setting physiologically interpretable thresholds on the extracted features.

Another part of this thesis focused on the video-based estimation of vital signs. The recently introduced iPPG technique only requires a digital video-camera and algorithms able to extract the subtle skin color changes induced by the cardiac pulse. After introducing the state-of-art methods for video-based HR estimation, different aspects were studied, aiming at providing a better understanding of this technology and at increasing its reliability. For this purpose, we used as databased acquired at CSEM on adult subjects. A small analysis was first performed to determine which parts of the face were the most relevant to compute HR. For this purpose, the subject faces were divided into 260 small regions of interest (ROIs) and a spectral analysis was performed for each region. We quantified the amount of power at the true HR and concluded that the forehead was the best region to derive HR and used it to derive the iPPG signals in the following analyses.

We also investigated the possibility to monitor the HR in the dark, by using a near-infrared camera and an infrared illumination. The results obtained were almost as good as for artificial light conditions using the RGB camera and we concluded that video-based HR monitoring in the dark is possible. Regarding HR estimation from the iPPG signal, we evaluated three different algorithms suitable for almost real-time frequency estimation, namely an algorithm based on sliding window adaptive singular value decomposition (SWASVD), a notch filter bank (NFB) and an adaptive frequency tracking algorithm (OSC-MSE-W). The OSC-MSE-W algorithm, which led to the smallest average absolute error, was selected for further analyses. Then, different existing iPPG derivation methods were implemented [96, 105, 106] (suitable for real-time application) and tested in addition to the basic pixel averaging method [5]. We concluded that it was more advantageous to combine the iPPG signals derived using different methods than using a single input. Even after increasing the robustness of the HR estimation by combining different iPPG signals and using an ROI tracking and skin segmentation algorithm, the results were not satisfactory for the *moving* subset of the database. For this reason, we developed a signal quality index, which was used to freeze the adaptation of the adaptive band-pass filter used for the estimation of HR and to label bad-quality segments for which the estimated HR was not reliable. These algorithmic developments resulted in HR estimation scheme, suitable for real-time applications. In addition to have a very small estimation delay (3-4 seconds), it is, to the best of our knowledge, the first approach which takes advantage of multiple iPPG signal derivation method.

This study was performed in the context of the Nano-Tera project NewbornCare, which final aim is to apply this technology to the monitoring of preterm infants in the neonatal intensive care unit. Preliminary results on neonates were presented in Appendix B. The results were encouraging, but we concluded that different steps of the processing need improvements in order to achieve a performance similar to the one achieved in controlled experimental conditions with adult subjects.

Some additional investigations were performed to study the iPPG-derived pulse rate variability (PRV). The database on adult subjects was used to assess the agreement between ECG-HRV and iPPG-PRV indexes. It was concluded that, in the non-stationary conditions induced by respiration and handgrip exercises, iPPG-PRV was not a good surrogate of ECG-HRV. We also investigated the feasibility to estimate the respiratory rate in real-time using the iPPG-IBI (in collaboration with Leila Mirmohamadsadeghi, former PhD student of ASPG). Promising results were obtained.

Different topics related to the processing of (video-based) photopleysmography were studied in this dissertation. The presented studies aimed at 1) developing signal processing tools to increase the reliability of the estimated HR and 2) assessing the potential of PPG/iPPG for other cardiovascular monitoring applications. Regarding the HR estimation, we concluded that adaptive frequency tracking tools have a real potential, both for conventional PPG and video-based iPPG signals. Indeed, due to their oscillatory behavior, these signals are perfect candidates for this type of methods. These methods allowed us to build robust HR estimation schemes with small estimation delays, which gives them an advantage over the classical block-based methods. We also noticed that, when dealing with PPG/iPPG signals, the derivation and the subsequent use of a signal quality index can significantly increase the accuracy of the estimated HR and is helpful to assess the veracity of detected events.

10.1 Summary of achievements

The major achievements of this dissertation, which were described in detail in the previous section, can be summarized as:

- Development of a processing scheme to perform robust HR estimation using wrist-type PPG signals during physical exercise.
- Reduction of false arrhythmia alarm rates in the ICU by the exploitation of signals from independent sources.
- Introduction of novel PPG-wave measures for the detection of AF using PPG signals.
- Development of a processing scheme to perform real-time HR estimation from video streams.
- Other contributions related to video-based monitoring:
 - Assessment of iPPG-derived pulse rate variability indexes in non-stationary conditions.
 - Real-time respiratory rate estimation using iPPG interbeat intervals.
 - Development of a signal quality index.

10.2 Perspectives

Algorithm implementation

Nowadays, people like to track their health during their daily activities and many wearable devices are available on the market for personal use. Forecasts regarding consumer health wearables are promising; this segment is expected to grow at compound annual growth rate of 28.3% (2016-2021) according to a survey by Frost and Sullivan. Among these health wearables, many of them have an embedded PPG sensor. There is therefore a real need for PPG processing algorithms to satisfy the consumer needs in this rapidly evolving field. The development of such processing tools was studied in this thesis. The next step to bring this research to market is the algorithm implementation. On-chip algorithm implementation is challenging as the aim is to achieve the lowest level of power consumption, while having accurate algorithms. Low-power signal processing embedded in wearable sensor nodes is beneficial as it results in reduced system power consumption, minimized system latency, increased device functionality (e.g. alarms can be triggered) [161].

This implementation aspect was kept in mind throughout the development of the different processing schemes and we believe that a future on-chip implementation is conceivable.

Regarding video-based monitoring, the development of a commercial monitor would also require an algorithm implementation step. In that case, the aim is to embark the algorithms for HR estimation, such that they run in real-time on the camera processor. The part related to the tracking of the ROI is expected to be most challenging, as it is the most computationally expensive.

A PPG-based cardiac monitor to screen for atrial fibrillation

AF is the most common cardiac arrhythmia, affecting 1-2% of the general population. AF can be asymptomatic in its initial course, which renders its diagnostic difficult. In the future, PPG-based cardiac monitors could be used to screen for AF in large population. It is therefore of interest to pursue the study of AF detection using PPG. We presented in this dissertation some

results computed from patients undergoing a cardiac ablation. It would be interesting to evaluate the performance of the proposed features on additional PPG data recorded in an ambulatory setting.

Further investigations into video-based monitoring on vital signs

This technology is quite recent and not yet fully understood. For example, more investigations should be performed to have a better understanding of the iPPG signal distortions induced by the different types of movements and illumination changes. It would be also interesting to see if adaptive noise cancellation techniques could be applied successfully to reconstruct clean iPPG signals during corrupted periods.

As mentioned in Appendix B, we believe that the tracking of the ROI could be improved by adding a dynamic feedback mechanism using the estimated HR and the signal quality index. More generally, efforts should be made to have a better synergy between the video processing and the signal processing parts.

The developed HR estimation scheme could be applied on many other applications. For example, it could be used to monitor drivers and detect drowsiness. It could also be used to monitor elderly people at home. Other applications such as the monitoring of people at customs and airports are also conceivable.

We showed in this dissertation that the iPPG PRV can be used to derive the respiratory rate. This is one possible application, but there are many other perspectives. A study demonstrated that the PRV derived from facial videos can be used to detect AF [151]. iPPG-derived cardiovascular parameters could also be used to detect some precursors signs of sudden infant death syndrome (higher basal HR, reduced HRV and lower parasympathetic activity and/or higher sympathovagal balance), as suggested in [162].

Appendix

Classifier Evaluation and Ensemble of Classifiers



A.1 Evaluation of classifiers¹

Machine learning is the study of models that can learn from input instances (known as training data), and subsequently make prediction on unseen instances (known as testing or prediction data). Having built a prediction model, commonly referred to as classifier, one can evaluate its performance by comparing classifier predictions with the actual testing data. In a two-class (known as binary) problem, the training data is divided into positive and negative classes. The positive class usually represents the instances for which the model is being created. For instance, in an HIV test, instances representing the HIV class are considered as members of the positive class whereas all other instances are considered as members of the negative class. Based on the predictions made by the classifier, instances can belong to one of the following categories:

- **True Positive (TP):** Instances that belong to the positive class and classified as such.
- **False Positive (FP);** known as the type I error, represent instances that belong to the positive class but classified as a member of the negative class.
- **True Negative (TN):** Instances that belong to the negative class and classified as such.
- **False Negative (FN);** known as the type II error, represent instances that belong to the negative class but classified as a member of the positive class.

A confusion matrix (also known as the contingency table) is usually used to visualize the performance of a specific classifier, as illustrated in Table A.1.

1. The content of this appendix is taken from the work of Sasan Yazdani (ASPG), 'Novel Low Complexity Biomedical Signal Processing Techniques for Online Applications', thesis director Dr. J.M. Vesin.

Table A.1: Confusion matrix of hypothetical model trained for a binary classification problem.

		Actual Class	
		+	-
Predicted Class	+	True Positive	False Positive Type I Error
	-	False Negative Type II Error	True negative

By closely studying the TP, FP, TN, and FN, one can evaluate the performance of a trained model. Over the years, several conventional metrics have been used in the literature, the most common ones are reviewed next.

- **Sensitivity (Se)**, also known as recall and the true positive rate (TPR), represents the number of instances that were classified as positive, which were indeed members of the positive class. In other words, this measure can be interpreted as the probability of a test to be positive given that the sample belong to the positive class. Se can therefore be calculated as,

$$\text{Sensitivity } (Se) = \frac{TP}{TP + FN} \quad (\text{A.1})$$

- **Specificity (SPC)**, also known as the true negative rate (TNR), is the proportion of instances that belong to the negative class and are classified as such. SPC calculates the same measure as Se, but for the negative class. SPC is calculated through,

$$\text{Specificity } (SPC) = \frac{TN}{TN + FP} \quad (\text{A.2})$$

- **Positive Prediction Value (PPV)**, also known as precision, studies how reliable the classifier is, in case the prediction is positive for an input sample. PPV is obtained as,

$$\text{Positive Prediction Value } (PPV) = \frac{TP}{TP + FP} \quad (\text{A.3})$$

- **Negative Prediction Value (NPV)** studies how reliable the classifier is, in case the prediction is negative for an input sample. NPV is obtained as follows:

$$\text{Negative Prediction Value } (NPV) = \frac{TN}{TN + FN} \quad (\text{A.4})$$

- **Accuracy (ACC)** is the fraction of test samples that are correctly predicted. ACC is measured as,

$$\text{Accuracy } (ACC) = \frac{TP + TN}{TP + FP + TN + FN} \quad (\text{A.5})$$

- **Detection Error Rate (DER)** represents the fraction of test samples that are incorrectly predicted. DER is defined as,

$$\text{Detection Error Rate } (DER) = 1 - ACC = \frac{FP + FN}{TP + FP + TN + FN} \quad (\text{A.6})$$

ACC and DER are simple statistics that report the overall performance of the trained classifier. Although these measures are intuitive and easy to calculate, they cannot provide reliable evaluation of the trained model in case the training data is unbalanced, i.e. one class has significantly more samples compared to the other. Let us assume a hypothetical machine learning scenario for which 95% of the dataset is comprised of the positive class (5% negative). If the

trained model classifies all input instances as members of the positive class, the final accuracy of the model is 0.95 (DER is 0.05). Therefore, one needs to study Se and PPV alongside ACC in order to perform a good performance evaluation. Alternatively, one can use the F-score (also known as F-measure), which evaluates the classifier by studying Se and PPV. F-score is calculated through,

$$F_{\beta} = (1 + \beta^2) \frac{PPV \cdot Se}{(\beta^2 \cdot PPV) + Se} \quad (A.7)$$

where β is the parameter that determines the relative weights of Se and PPV. It is common to compute the F-score with $\beta = 1$, which represents the balanced F-score. This harmonic mean between Se and PPV is known as the F1-score,

$$F_1 = 2 \frac{PPV \cdot Se}{PPV + Se} \quad (A.8)$$

Over the years, several other metrics such as the Phi and Matthews correlation coefficients have been proposed in the literature in order to take other important factors such as the true negative rate into consideration. Further information on these metrics can be found in [163, 164]. In the context of this dissertation, only two-class machine learning techniques are studied, methods described in this appendix covers binary classification evaluation. More information on multi-class performance evaluation can be found in [164?].

A.2 Classifier combination: selection and fusion

Over the pas decades, the idea of training multiple classifiers and combining them to obtain the desired output has gained momentum. Multi expert (classifier) techniques can be categorized into two sub-branches, namely selection and fusion. Consider a hypothetical problem in which some classifiers outperform others at specific tasks, but their overall performance does not. For instance, imagine a classification problem in which the goal is to identify cats and dogs, and that we have created two classifiers, C_1 and C_2 , from the training set. Assume that C_1 outperforms C_2 but does not perform well for small-sized cats and dogs. On the other hand, C_2 perfectly separates small cats and dogs but does poorly otherwise. In this case when C_2 is used for small-sized inputs while C_1 is used otherwise, the overall prediction performance improves. In this way the classification task is broken into two subspaces, where C_1 performs better in one while C_2 in the other. This combination technique is referred to as classification selection. Figure A.1 illustrates an example fusion technique with four classifiers.

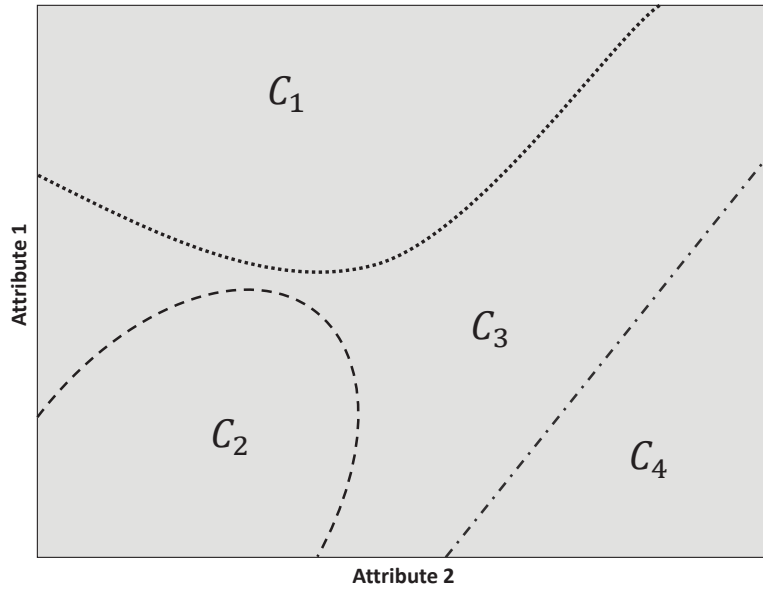


Figure A.1: An example classifier selection technique illustrated in a two attribute subspace. All four classifiers (C_1 , C_2 , C_3 , and C_4) work in coherence, each acting as the expert in its respective subspace.

Another technique to combine classifiers is through classifier fusion. In this scheme, n classifiers with low bias and high variance are created separately on the training set and their respective predictions are combined to obtain the final prediction [164]. Figure A.2 illustrates the framework used to create typical fusion of classifiers.

Various voting techniques can be used in classifier fusion, details of which are described as follows.

- **Majority Voting** is the simplest yet most common voting scheme. As suggested by its name, the prediction label in this scheme is selected as the label that has been predicted by majority of the trained classifiers, i.e statistical mode of the prediction labels.
- **Weighted Majority Voting** works similarly to majority voting, however the predictions made by classifiers are weighted, with some classifiers (the better performing and more complex ones) having greater influence in label prediction.
- **Optimistic Voting**. In this scheme the final prediction label is considered as positive, if at least one of the trained classifiers prediction is as such.
- **Pessimistic Voting**, also known as unanimous voting, is an scheme in which the final prediction label is considered as positive, if the trained classifiers unanimously predict as such.

A.3 The wisdom of crowds

The wisdom of crowds [165, 166] is a theory suggesting that, given the right conditions, large groups of average people are smarter than a few elite experts. There are four necessary

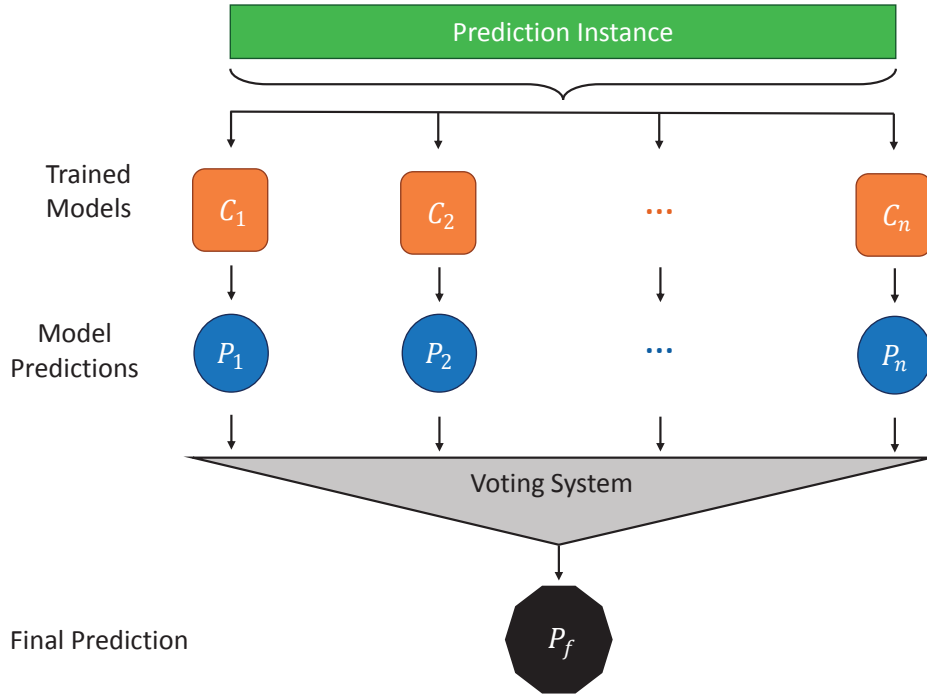


Figure A.2: Typical framework of classifier fusion. In order to obtain the final prediction (P_f) for an input instance, the predicted label (P_1, P_2, \dots, P_n) for each trained classifiers (C_1, C_2, \dots, C_n) taken into account.

conditions for a crowd to be considered as wise. First, the rule of independence, which states people's opinion should not be influenced by others. Second, the rule of decentralization, stating people should be able to draw their opinion based on their expertise. Third, the diversity of opinion, meaning people should have some information on the problem. Finally, the aggregation rule, which is a scheme to combine the crowd's opinion to obtain a final decision.

Implementation of the wisdom of crowds in machine learning is somewhat difficult, even though the aggregation and independence rules are naturally implemented when combining classifiers. Decentralization is also not hard to implement. This rule states that members of an ensemble of classifiers should perform prediction based on different specialties (expertise). Although meta algorithms such as bagging, boosting, rotation forest, and etc. can have numerous base classifiers, in their implementation all base classifiers are essentially the same and are only exposed to different training data [164]. Nevertheless, in machine learning, decentralization can be carried out by training predictors of different natures, for instance a support vector machine, a decision tree, along side a multi-layer perceptron neural network. Furthermore, if accurate enough classifiers are created, one does not need numerous models to reach reliable predictions. The difficulty in the implementation of wisdom of crowd lies in the diversity of opinion. When creating classifiers, several attributes (features) are extracted from the training data. for the diversity rule to carry out, classifiers need to be trained on different features which in practice does not make sense as they can improve the performance of other train classifiers. Moreover, often features can be unintuitive to extract and therefore, are defined by the expert in the field of the study.

The implementation of wisdom of crowds, however, is not impossible. As pointed out Section 6.4, during the false alarm detection Physionet/CinC challenge, the organizers took the top performing algorithms and created a meta-algorithm, which obtained an accuracy higher than that of any individual algorithm. The diversity of opinion rule was perfectly met as different challengers used different features to train their classifiers. This can be considered as an example implementation of wisdom of crowds, as all four rules were applied.

Video-Based Heart Rate Monitoring in Neonatal Intensive Care Unit: Preliminary Results

B

B.1 Introduction

In Switzerland, approximately 800 infants are born preterm every year, with a gestational age below 32 weeks [167]. This phenomenon is currently growing. Indeed, over the last three decades, the number of children born with a birth weight below 1500g has doubled [168]. Different factors such as changing demographics, older maternal age and higher rates of medically assisted reproductions are related to this phenomenon [169]. Preterm infants, who require special care, are first placed in a Neonatal Intensive Care Unit (NICU), where their vital sign are continuously monitored. More specifically, their heart rate (HR), respiratory rate and oxygen saturation are measured. The ECG monitor used to derive HR requires the attachment of adhesive electrodes to the fragile skin of these infants, which induces a discomfort. Moreover, the current monitoring systems are much prone high false alarm rates [76]. In addition, the brain, which is the most sensitive organ, is not monitored yet. The aim of the NewbornCare^{1 2} project is to overcome the limitations of the current NICU monitoring systems. It has two main components; the multi-sensor component, which aims at monitoring arterial and cerebral oxygen saturation and the central component, which aims at monitoring respiratory rate and HR in a contactless manner. The latter, discussed in this appendix, is achieved by the means of a vision system consisting of two cameras and an infrared illumination.

Aarts *et al.* investigated for the first time the feasibility of continuous non-contact HR monitoring in the NICU [170]. In this study, a total of 19 infants were monitored for a duration between one and five minutes. The iPPG signal, obtained by the averaging of the green channel pixel values, was used to compute a joint-time-frequency diagram. A good match, defined as $|\text{HR}_{ref} - \text{HR}_{est}| < 5$ bpm more than 90% of the time, was obtained for 13 sequences out of the 19 sequences. Different challenges were encountered and the authors concluded that algorithms improvements are necessary to achieve more reliable results.

Villarroel *et al.* also evaluated the feasibility of continuous non-contact monitoring of cardiorespiratory vital signs in the NICU [171]. In this study, two infants were monitored for 39.8 hours. Among these 39.8 hours, 20.1 hours of stable periods of video data were first identified and la-

1. The NewbornCare RTD project was funded by NanoTera.

2. In the news: <https://actu.epfl.ch/news/medically-monitoring-premature-babies-with-cameras/>

beled suitable for HR estimation. ICA was applied to the raw RGB traces and a band-pass filter with cut-off frequencies of 78 and 300 bpm was then used. An auto-regressive model was used to identify the cardiac frequency from the pre-processed iPPG signals, based on the methodology proposed in [125] by the same authors. An 8-second sliding window was used for this purpose. The resulting average absolute error was 2.83 bpm.

The aim of this appendix is to present the preliminary results obtained on a database acquired at the UniversitätsSpital Zürich (USZ) NICU³. For this purpose, the real-time HR estimation scheme presented in Chapter 8 was adapted to capture any HR value within the HR range of preterm infants.

B.2 Methods

B.2.1 Database

The NIR and RGB cameras were positioned on the side of the incubator/bed, as shown in Figure B.1 (the same cameras as for the adult database were used, see Section 8.2). The duration of each sequence was 20 minutes (40 minutes for some sequences). The reference HR, as well as the ECG signal, were given by the NICU Däger commercial monitor. These ground-truth data were synchronized with the data from the vision system⁴. The results presented in this appendix were generated on a subset of the entire database. This subset consists of 13 sequences acquired on 10 different preterm infants, corresponding to a total of five hours of recording. This database subset is summarized in Table B.1. The name and duration of each sequence are indicated in this table, as well as the level of activity of the baby during the recording, the illumination conditions and some specific remarks to depict the challenges encountered in the NICU.

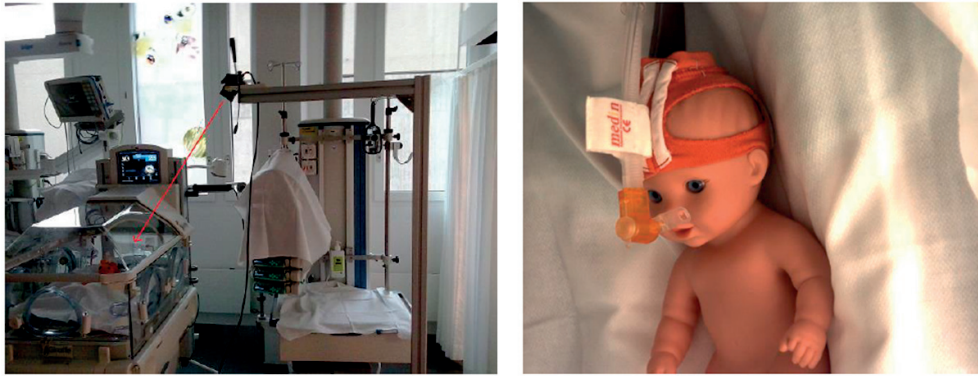


Figure B.1: Experimental setup used for the database acquisition at the USZ NICU. Left: camera setup. Right: example of an image obtained with this camera setup.

B.2.2 Tracking and skin segmentation

Tracking of the selected ROI and skin segmentation were performed with the method already described in Section 8.3.1. For most of the sequences, the ROI was selected on the baby's face. However, the area of visible skin was sometimes very restricted due to the presence of medical

3. The acquisition of the video sequences was supervised by Tanja Karen (USZ), Virginie Moser (CSEM) and Daniel Ostojic (USZ).

4. Data synchronization was done by Virginie Moser (CSEM) and Fabian Braun (CSEM).

Table B.1: The neonate database subset analyzed in this appendix. Level of activity: subjective measure of baby’s level of activity during the recording; Illumination conditions: subjective assessment of quality of the illumination conditions (poor indicates typically a very dark environment or an illumination continuously changing); Remarks: events susceptible to impact the contactless measure of HR.

Sequence name	Duration [min]	Inside incubator?	Level of activity	Illumination conditions	Remarks
P3-R1	20	no	high	poor	-
P6-R1	20	no	low	moderate	-
P6-R3	20	no	high	good	eating
P7-R1	20	no	low	good	many sensors
P7-R3	20	no	low	good	-
P11-R1	20	no	moderate	moderate	pacifier
P11-R3	20	no	moderate	good	-
P12-R1	20	no	high	moderate	-
P13-R1	20	no	low	poor	-
P15-R1	40	no	very high	good	many occlusions
P19-R1	20	yes	low	good	-
P20-R1	20	yes	moderate	moderate	light reflection (plexiglas)
P21-R1	40	yes	low	moderate	ROI on the back

equipment (feeding tube, continuous positive airway pressure, oxygen mask, etc.). The size and position of the selected ROI were therefore changing from one recording to another.

Regarding the skin segmentation step, it was difficult to find a threshold value working well for all patients, given the very variable lightening conditions and the different skin tones. Therefore, for each patient, the first step consisted of the manual selection of a patch of skin on the first frame of the video, which was then used to apply automatic skin segmentation for the whole duration of the recording.

B.2.3 iPPG signal derivation and heart rate estimation

The HR estimation scheme presented in Chapter 8 was adapted for this application. As shown in section Section 8.3.4, different iPPG signal derivation techniques were used to compute the following signals:

- $iPPG_{\text{green}}$.
- $iPPG_{\text{hue}}$.
- $iPPG_{\text{SSR}}$, using $l = 15$ samples.
- $iPPG_{\text{POS}}$, using $l = 15$ samples.
- $iPPG_{\text{GRD}}$.

However, it was not possible to use the same band-pass filter as for adults because the babies have a higher HR. According to [172], “The minimum heart rate of premature or low-birth-weight infants can be as low as 73 bpm; the maximum heart rate can be as high as 211 bpm.”. An FIR band-pass filter with cut-off frequencies of 54 and 222 bpm was therefore used to compute $iPPG_{\text{green}}$, $iPPG_{\text{hue}}$ and $iPPG_{\text{GRD}}$ signals. The $iPPG_{\text{SSR}}$ and $iPPG_{\text{POS}}$ signals were also band-pass filtered using the same filter.

After visual inspection of the derived iPPG signals, it was noticed that, in these realistic experimental conditions, the $iPPG_{\text{green}}$ and $iPPG_{\text{hue}}$ signals were usually of bad quality and the largest peak on the PSD was not always at the HR frequency, as it was usually the case in the controlled

experimental conditions of the adult database. This is illustrated in Figure B.2. In this figure, it can be observed that for the studied 8-second window, the main peak on the PSD was at the HR frequency for iPPG_{SSR}, iPPG_{POS} and iPPG_{GRD} signals, but it not for the iPPG_{green} and iPPG_{hue} signals. For this reason, the iPPG_{green} and iPPG_{hue} signals were discarded.

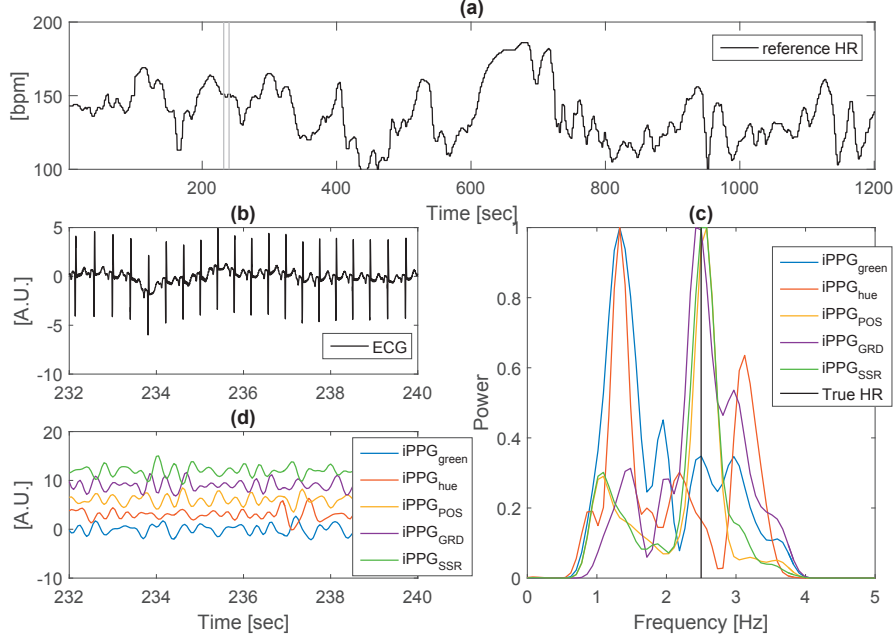


Figure B.2: Example of iPPG signals and their PSD for an 8-second data segment (sequence P3-R1). (a) reference HR and selected time window; (b) ECG signal; (c) normalized PSD of the iPPG signals and true HR; (d) iPPG signals.

The quality of the iPPG signals was then evaluated. For this purpose, the signal quality index (SQI) was computed using the method proposed in Section 8.3.5. Then, the modified version of the multi-input adaptive frequency tracking algorithm (OSE-MSE-SQI) described in Section 2.3.3 was used to estimate the HR. The iPPG_{SSR}, iPPG_{POS} and iPPG_{GRD} signals and their respective SQIs were provided as inputs. The HR estimates for which at least two of the iPPG signals were bad quality ($SQI_m[n] = 0$) were considered as not reliable and not taken into account in the error calculation.

B.3 Results

After removing the bad-quality segments automatically identified using the SQI, an HR estimate was available 92.8% of the time. The average absolute error (AAE) between the estimated HR and the reference HR was computed on a 5-second sliding window (4-second overlap). The percentage of samples for which the absolute difference between the estimated and the reference HR ($|HR_{ref}[n] - HR_{est}[n]|$) was lower than 5 bpm was also calculated and reported. The results, for each sequence and globally, are shown in Table B.2. Figures B.3, B.4 and B.5 show the HR estimates for three sequences. The estimated HR alongside the reference HR is shown on the top. The grey zones correspond to regions where at least one the the signal was bad-quality ($SQI_m[n] = 0$). The DI measure of inter-frame variability is shown on the bottom plot.

This measure, which was used in the computation of the SQI, gives an idea about the amount of disturbances.

Table B.2: Results. For each sequence, the percentage of removed data, the AAE and the percentage of time the absolute error was lower than five bpm are reported.

Sequence name	Duration [min]	removed data [%]	AAE [bpm]	$ HR_{ref} - HR_{est} < 5 \text{ bpm}$ [%] of the time
P3-R1	20	6.6	14.5	43.3
P6-R1	20	4.4	2.4	91.2
P6-R3	20	3.4	6.6	74.9
P7-R1	20	4.3	3.8	78.4
P7-R3	20	1.3	1.7	94.8
P11-R1	20	18.9	12.6	45.6
P11-R3	20	10.8	6.4	82.6
P12-R1	20	10.3	6.1	73.8
P13-R1	20	2.1	11.1	61.6
P15-R1	40	21.3	5.7	79.3
P19-R1	20	0.5	5.7	81.2
P20-R1	20	0.9	7.1	73.4
P21-R1	40	1.3	1.9	97.1
Overall	300	7.2	6.11 + 12.93	77.4

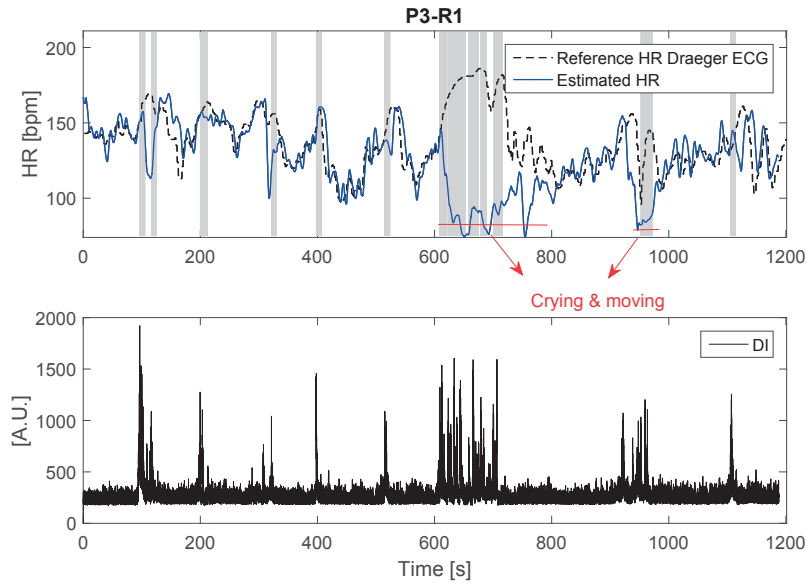


Figure B.3: Sequence P3-R1. Estimated and reference HRs (top) and DI measure of inter-frame variability (bottom).

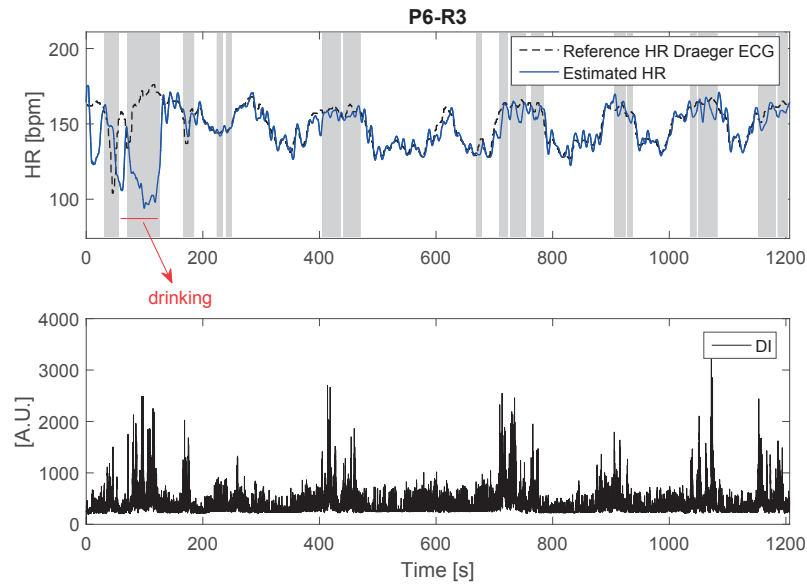


Figure B.4: Sequence P6-R3. Estimated and reference HRs (top) and DI measure of inter-frame variability (bottom).

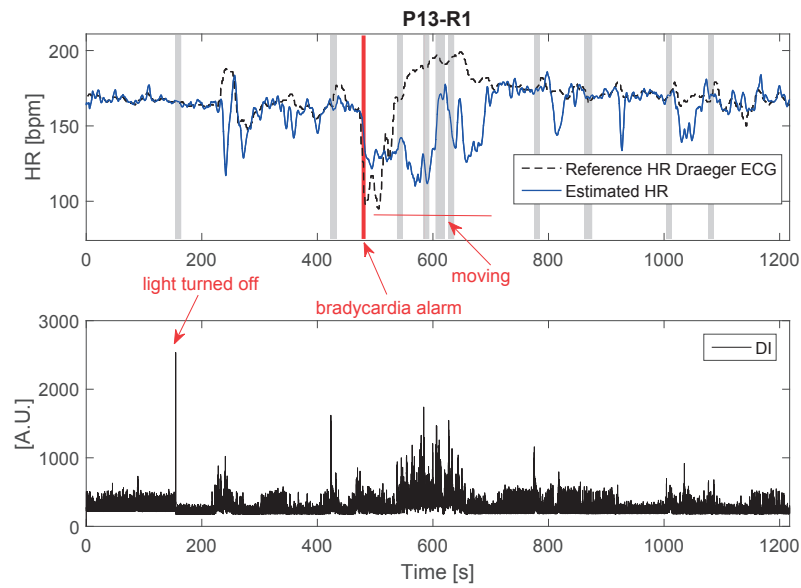


Figure B.5: Sequence P13-R1. Estimated and reference HRs (top) and DI measure of inter-frame variability (bottom).

B.4 Discussion and conclusion

These preliminary results are encouraging. According to the proposed SQI, an HR estimate was available 92.8% of the time. During this time, the overall AAE was 6.11 bpm and the HR estimate was very accurate 77.4% of the time. It can be noticed in Table B.2 that the results were quite variable from one sequence to another. The worst results were obtained for the sequence P3-R1 (AAE of 14.5 bpm), which is shown in Figure B.3. In this figure, it can be seen that the estimated HR follows well the reference HR, except during two long periods. The baby was crying and moving a lot during these periods, resulting in numerous fast movements difficult to track for the ROI tracking algorithm. In addition, as specified in Table B.1, the illumination conditions were really poor for this recording. Different factors have therefore contributed to the deterioration of the computed iPPG signals. On the other hand, the example shown in Figure B.4 is very encouraging for multiple reasons. Firstly, this infant had a dark skin tone, suggesting that the proposed processing scheme was not affected by this factor. Secondly, this baby was continuously active (gentle movements) during the whole duration of the recording. Finally, the HR fluctuated a lot. Despite of all these challenges, the estimated HR matches really well the reference HR (AAE of 3.4 bpm).

The example depicted in Figure B.5 shows an a recording during which a bradycardia occurred. The estimated HR decreases during this event, but not as much as expected because the baby started to move right after the event. These movements, combined with the poor illumination conditions, resulted in an inaccurate HR estimate for a few minutes after the alarm.

More generally, it can be observed that the results presented in Table B.2 can be correlated with the level of activity and the illumination conditions reported in Table B.1. Bad illumination conditions and high activity level seem to affect the performance of HR estimation, which is not really surprising. However, performing the estimation from a very restricted surface of skin does not seem detrimental (see for example sequence P7-R1), provided that the automatic skin segmentation is working well. The variation in the baby's activity patterns was described as a challenge by Villarroel *et al.* [171]. Our observations also showed that preterm infants make irregular movements, even when they sleep, which can impact the remote measure of HR.

As mentioned already, these results are preliminary results and different aspects of the processing need more investigations in order to increase the reliability of the estimated HR in these challenging conditions. We believe that the following points should be examined in more detail in future work:

- Motion artifacts: further work should focus on the improvement of the robustness to motion. This implies some improvements of the ROI tracking algorithm. For example, several ROIs could be tracked simultaneously. Then, a dynamic feedback mechanism could be added, in which the developed iPPG SQI could be used to select the most reliable ROI.
- Monitoring in the dark: it will be interesting to evaluate the performance of the HR estimation using the NIR camera in the dark. Given that we use only one method to derive the iPPG signals in that case, the HR estimate might be less accurate.
- Use of the respiratory information to increase the robustness of the HR estimation scheme: the respiratory information could be used to improve HR estimation. For example, we noticed that, in some cases, in the PSD of the iPPG signal, the peak corresponding to the first harmonic component of the respiration was larger than the peak corresponding to the HR. Information about the instantaneous respiration rate could therefore be useful to make an adaptive cancellation of the first harmonic component of the respiration in such cases. For this purpose, the algorithm developed in the context of this NewbornCare project to

estimate the respiration from the thoracic motion could be used [173].

- Detection of life-threatening events: as the aim is to replace the conventional sensors by the camera system in the future, it will be necessary to report not only the accuracy of the estimated parameters, but also the sensitivity obtained for the detection of HR-related alarms. Indeed, for these life-threatening events, only a system with a sensitivity of 100% can be conceivable to replace the traditional systems.

To conclude, iPPG is a promising technology, but algorithmic improvements are still required to deal with various challenges and to have systems that are clinically reliable to monitor fragile infants in the NICU. It should be mentioned that video-based monitoring of infants vital signs is also interesting for home monitoring applications. Indeed, it has been suggested recently that camera-based, non-contact vital signs monitoring technology may provide a way for early prevention of sudden infant death syndrome [162].

Bibliography

- [1] S. C. Mukhopadhyay, “Wearable Sensors for Human Activity Monitoring: A Review,” *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1321–1330, Mar. 2015.
- [2] H. Asada, P. Shaltis, A. Reisner, S. Rhee, and R. Hutchinson, “Mobile monitoring with wearable photoplethysmographic biosensors,” *IEEE Eng. Med. Biol. Mag.*, vol. 22, no. 3, pp. 28–40, May 2003.
- [3] Y. L. Zheng, X. R. Ding, C. C. Y. Poon, B. P. L. Lo, H. Zhang, X. L. Zhou, G. Z. Yang, N. Zhao, and Y. T. Zhang, “Unobtrusive Sensing and Wearable Devices for Health Informatics,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1538–1554, May 2014.
- [4] Y. Sun and N. Thakor, “Photoplethysmography Revisited: From Contact to Noncontact, From Point to Imaging,” *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 463–477, Mar. 2016.
- [5] W. Verkruijsse, L. O. Svaasand, and J. S. Nelson, “Remote plethysmographic imaging using ambient light,” *Opt. Express*, vol. 16, no. 26, pp. 21 434–21 445, Dec. 2008.
- [6] J. Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiol. Meas.*, vol. 28, no. 3, pp. R1–39, Mar. 2007.
- [7] H. Banaee, M. U. Ahmed, and A. Loutfi, “Data Mining for Wearable Sensors in Health Monitoring Systems: A Review of Recent Trends and Challenges,” *Sensors*, vol. 13, no. 12, pp. 17 472–17 500, Dec. 2013.
- [8] H. Haken, H. P. Koepchen, and H. Haken, Eds., *Rhythms in Physiological Systems*, ser. Springer Series in Synergetics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1991, vol. 55.
- [9] L. Sörnmo and P. Laguna, *Bioelectrical signal processing in cardiac and neurological applications*. Academic Press, 2005.
- [10] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-time signal processing*, 2nd ed. Upper Saddle River, N.J: Prentice Hall, 1999.
- [11] L. Griffiths, “Rapid measurement of digital instantaneous frequency,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 23, no. 2, pp. 207–222, Apr. 1975.
- [12] B. Widrow, J. R. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, J. E. Dong, and R. C. Goodlin, “Adaptive noise cancelling: Principles and applications,” *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692–1716, Dec. 1975.
- [13] J. Zeidler, E. Satorius, D. Chabries, and H. Wexler, “Adaptive enhancement of multiple sinusoids in uncorrelated noise,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 26, no. 3, pp. 240–254, Jun. 1978.

- [14] P. Regalia, "A Complex Adaptive Notch Filter," *IEEE Signal Process. Lett.*, vol. 17, no. 11, pp. 937–940, Nov. 2010.
- [15] N. I. Cho and S. U. Lee, "On the adaptive lattice notch filter for the detection of sinusoids," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 40, no. 7, pp. 405–416, Jul. 1993.
- [16] K. W. Martin and M. Padmanabhan, "Using an IIR adaptive filter bank to analyze short data segments of noisy sinusoids," *IEEE Trans. Signal Process.*, vol. 41, no. 8, pp. 2583–2590, Aug. 1993.
- [17] M. R. Petraglia, S. K. Mitra, and J. Szczupak, "Adaptive sinusoid detection using IIR notch filters and multirate techniques," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 41, no. 11, pp. 709–717, Nov. 1994.
- [18] T. Kwan and K. Martin, "Adaptive detection and enhancement of multiple sinusoids using a cascade IIR filter," *IEEE Trans. Circuits Syst.*, vol. 36, no. 7, pp. 937–947, Jul. 1989.
- [19] H.-E. Liao, "Two discrete oscillator based adaptive notch filters (OSC ANFs) for noisy sinusoids," *IEEE Trans. Signal. Process.*, vol. 53, no. 2, pp. 528–538, Feb. 2005.
- [20] Y. Prudat and J. M. Vesin, "Multi-signal extension of adaptive frequency tracking algorithms," *Signal Process.*, vol. 89, no. 6, pp. 963–973, Jun. 2009.
- [21] J. Van Zaen, L. Uldry, C. Duchêne, Y. Prudat, R. A. Meuli, M. M. Murray, and J.-M. Vesin, "Adaptive tracking of EEG oscillations," *J. Neurosci. Methods*, vol. 186, no. 1, pp. 97–106, Jan. 2010.
- [22] Y. Prudat, "Adaptive frequency tracking and application to biomedical signals," Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne, 2009.
- [23] J. Van Zaen, "Efficient schemes for adaptive frequency tracking and their relevance for EEG and ECG," Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne, 2012.
- [24] L. Uldry, C. Duchêne, Y. Prudat, M.-M. Murray, and J.-M. Vesin, "Adaptive tracking of EEG frequency components," in *Advanced Biosignal Processing*. Springer, 2009, pp. 123–144.
- [25] I. I. Goncharova and J. S. Barlow, "Changes in EEG mean frequency and spectral purity during spontaneous alpha blocking," *Electroencephalogr. Clin. Neurophysiol.*, vol. 76, no. 3, pp. 197–204, Sep. 1990.
- [26] H. Ugnell and P. A. Öberg, "The time-variable photoplethysmographic signal; dependence of the heart synchronous signal on wavelength and sample volume," *Med. Eng. Phys.*, vol. 17, no. 8, pp. 571–578, Dec. 1995.
- [27] S. Takatani and M. D. Graham, "Theoretical analysis of diffuse reflectance from a two-layer tissue model," *IEEE Trans. Biomed. Eng.*, vol. 26, no. 12, pp. 656–664, Dec. 1979.
- [28] T. Tamura, Y. Maeda, M. Sekine, and M. Yoshida, "Wearable photoplethysmographic sensors—Past and present," *Electronics*, vol. 3, no. 2, pp. 282–302, Apr. 2014.
- [29] Y. Maeda, M. Sekine, and T. Tamura, "The advantages of wearable green reflected photoplethysmography," *J. Med. Syst.*, vol. 35, no. 5, pp. 829–834, Oct. 2011.

- [30] J. Lee, K. Matsumura, K. Yamakoshi, P. Rolfe, S. Tanaka, and T. Yamakoshi, "Comparison between red, green and blue light reflection photoplethysmography for heart rate monitoring during motion," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2013, pp. 1724–1727, Jul. 2013.
- [31] J. Allen and A. Murray, "Age-related changes in the characteristics of the photoplethysmographic pulse shape at various body sites," *Physiol. Meas.*, vol. 24, no. 2, p. 297, Mar. 2003.
- [32] —, "Similarity in bilateral photoplethysmographic peripheral pulse wave characteristics at the ears, thumbs and toes," *Physiol. Meas.*, vol. 21, no. 3, p. 369, May 2000.
- [33] A. C. M. Dassel, R. Graaff, A. Meijer, W. G. Zijlstra, and J. G. Aarnoudse, "Reflectance pulse oximetry at the forehead of newborns: The influence of varying pressure on the probe," *J. Clin. Monit.*, vol. 12, no. 6, pp. 421–428, Nov. 1996.
- [34] X. F. Teng and Y. T. Zhang, "The effect of contacting force on photoplethysmographic signals," *Physiol. Meas.*, vol. 25, no. 5, pp. 1323–1335, Oct. 2004.
- [35] —, "The effect of applied sensor contact force on pulse transit time," *Physiol. Meas.*, vol. 27, no. 8, pp. 675–684, Aug. 2006.
- [36] A. Grabovskis, Z. Marcinkevics, U. Rubins, and E. Kviesis-Kipge, "Effect of probe contact pressure on the photoplethysmographic assessment of conduit artery stiffness," *J. Biomed. Opt.*, vol. 18, no. 2, p. 27004, Feb. 2013.
- [37] Y.-S. Yan, C. C. Poon, and Y.-T. Zhang, "Reduction of motion artifact in pulse oximetry by smoothed pseudo wigner-ville distribution," *J. Neuroengineering Rehabil.*, vol. 2, p. 3, Mar. 2005.
- [38] R. Krishnan, B. Natarajan, and S. Warren, "Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 8, pp. 1867–1876, Aug. 2010.
- [39] K.-A. Reddy, B. George, and V.-J. Kumar, "Use of fourier series analysis for motion artifact reduction and data compression of photoplethysmographic signals," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 5, pp. 1706–1711, May 2009.
- [40] G. Comtois and Y. Mendelson, "A noise reference input to an adaptive filter algorithm for signal processing in a wearable pulse oximeter," in *Bioengineering Conference IEEE 33rd Annual Northeast*, 2007, pp. 106–107.
- [41] M.-R. Ram, K.-V. Madhav, E.-H. Krishna, N.-R. Komalla, and K.-A. Reddy, "A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 5, pp. 1445–1457, May 2012.
- [42] K. Naraharisetti, M. Bawa, and M. Tahernezehadi, "Comparison of different signal processing methods for reducing artifacts from photoplethysmograph signal," in *2011 IEEE International Conference on Electro/Information Technology (EIT)*, 2011, pp. 1–8.
- [43] Z. Zhang, Z. Pi, and B. Liu, "TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise," *IEEE Trans. Biomed. Eng.*, vol. 62, pp. 522–531, Feb. 2015.
- [44] Z. Zhang, "Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 1902–1910, Aug. 2015.

- [45] N. Lakshminarasimha Murthy, P. Madhusudana, P. Suresha, V. Periyasamy, and P. Ghosh, "Multiple spectral peak tracking for heart rate monitoring from photoplethysmography signal during intensive physical exercise," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2391–2395, Dec. 2015.
- [46] E. Khan, F. A. Hossain, S. Z. Uddin, S. K. Alam, and M. K. Hasan, "A robust heart rate monitoring scheme using photoplethysmographic signals corrupted by intense motion artifacts," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 550–562, Mar. 2016.
- [47] B. Sun and Z. Zhang, "Photoplethysmography-based heart rate monitoring using asymmetric least squares spectrum subtraction and bayesian decision theory," *IEEE Sensors J.*, vol. 15, no. 12, pp. 7161–7168, Dec. 2015.
- [48] S. Fallet and J. M. Vesin, "Adaptive frequency tracking for robust heart rate estimation using wrist-type photoplethysmographic signals during physical exercise," in *Computing in Cardiology*, 2015, pp. 925–928.
- [49] S. Fallet and J.-M. Vesin, "Robust heart rate estimation using wrist-type photoplethysmographic signals during physical exercise: an approach based on adaptive filtering," *Physiol. Meas.*, vol. 38, no. 2, pp. 155–170, Feb. 2017.
- [50] P. S. R. Diniz, *Adaptive Filtering*. Springer US, 2013.
- [51] B. Widrow and S. D. Stearns, *Adaptive signal processing*, ser. Prentice-Hall signal processing series. Englewood Cliffs, N.J: Prentice-Hall, 1985.
- [52] K. Chan and Y. Zhang, "Adaptive reduction of motion artifact from photoplethysmographic recordings using a variable step-size LMS filter," in *Proceedings of IEEE Sensors*, 2002, vol. 2, 2002, pp. 1343–1346 vol.2.
- [53] J. Allen and A. Murray, "Age-related changes in peripheral pulse timing characteristics at the ears, fingers and toes," *J. Hum. Hypertens.*, vol. 16, no. 10, pp. 711–717, Oct. 2002.
- [54] M. Chan, D. Estève, J.-Y. Fourniols, C. Escriba, and E. Campo, "Smart wearable systems: Current status and future challenges," *Artif. Intell. Med.*, vol. 56, no. 3, pp. 137–156, Nov. 2012.
- [55] V. Fuster, L. E. Rydén, R. W. Asinger, D. S. Cannom, H. J. Crijns, R. L. Frye, J. L. Halperin, G. N. Kay, W. W. Klein, S. Lévy, R. L. McNamara, E. N. Prystowsky, L. S. Wann, D. G. Wyse, R. J. Gibbons, E. M. Antman, J. S. Alpert, D. P. Faxon, V. Fuster, G. Gregoratos, L. F. Hiratzka, A. K. Jacobs, R. O. Russell, S. C. Smith, W. W. Klein, A. Alonso-Garcia, C. Blomström-Lundqvist, G. de Backer, M. Flather, J. Hradec, A. Oto, A. Parkhomenko, S. Silber, A. Torbicki, American College of Cardiology/American Heart Association Task Force on Practice Guidelines, European Society of Cardiology Committee for Practice Guidelines and Policy Conferences (Committee to Develop Guidelines for the Management of Patients With Atrial Fibrillation), and North American Society of Pacing and Electrophysiology, "ACC/AHA/ESC Guidelines for the Management of Patients With Atrial Fibrillation: Executive Summary A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines and Policy Conferences (Committee to Develop Guidelines for the Management of Patients With Atrial Fibrillation) Developed in Collaboration With the North American Society of Pacing and Electrophysiology," *Circulation*, vol. 104, no. 17, pp. 2118–2150, Oct. 2001.

- [56] European Heart Rhythm Association, European Association for Cardio-Thoracic Surgery, A. J. Camm, P. Kirchhof, G. Y. H. Lip, U. Schotten, I. Savelieva, S. Ernst, I. C. V. Gelder, N. Al-Attar, G. Hindricks, B. Prendergast, H. Heidbuchel, O. Alfieri, A. Angelini, D. Atar, P. Colonna, R. D. Caterina, J. D. Sutter, A. Goette, B. Gorenek, M. Heldal, S. H. Hohloser, P. Kolh, J.-Y. L. Heuzey, P. Ponikowski, F. H. Rutten, E. C. f. P. G. (cpg), A. Vahanian, A. Auricchio, J. Bax, C. Ceconi, V. Dean, G. Filippatos, C. Funck-Brentano, R. Hobbs, P. Kearney, T. McDonagh, B. A. Popescu, Z. Reiner, U. Sechtem, P. A. Sirnes, M. Tendera, P. E. Vardas, P. Widimsky, D. Reviewers, P. E. Vardas, V. Agladze, E. Aliot, T. Balabanski, C. Blomstrom-Lundqvist, A. Capucci, H. Crijns, B. Dahlöf, T. Folliguet, M. Glikson, M. Goethals, D. C. Gulba, S. Y. Ho, R. J. M. Klautz, S. Kose, J. McMurray, P. P. Filardi, P. Raatikainen, M. J. Salvador, M. J. Schali, A. Shpektor, J. Sousa, J. Stepinska, H. Ueetoo, J. L. Zamorano, and I. Zupan, "Guidelines for the management of atrial fibrillation," *Eur. Heart J.*, vol. 31, no. 19, pp. 2369–2429, Oct. 2010.
- [57] S. J. Podd, C. Sugihara, S. S. Furniss, and N. Sulke, "Are implantable cardiac monitors the 'gold standard' for atrial fibrillation detection? A prospective randomized trial comparing atrial fibrillation monitoring using implantable cardiac monitors and DDDR permanent pacemakers in post atrial fibrillation ablation patients," *Europace*, vol. 18, no. 7, pp. 1000–1005, Jul. 2016.
- [58] D. D. McManus, J. Lee, O. Maitas, N. Esa, R. Pidikiti, A. Carlucci, J. Harrington, E. Mick, and K. H. Chon, "A novel application for the detection of an irregular pulse using an iPhone 4s in patients with atrial fibrillation," *Heart Rhythm*, vol. 10, no. 3, pp. 315–319, Mar. 2013.
- [59] L. Krivoshei, S. Weber, T. Burkard, A. Maseli, N. Brasier, M. Kühne, D. Conen, T. Huebner, A. Seeck, and J. Eckstein, "Smart detection of atrial fibrillation," *Europace*, pp. 753–757, Jul. 2016.
- [60] A. G. Bonomi, F. Schipper, L. Eerikäinen, J. Margarito, R. M. Aarts, S. Babaeizadeh, H. M. Morree, and L. Dekker, "Atrial fibrillation detection using photo-plethysmography and acceleration data at the wrist," in *Computing in Cardiology*, Sep. 2016, pp. 277–280.
- [61] V. D. A. Corino, R. Laureanti, L. Ferranti, G. Scarpini, F. Lombardi, and L. T. Mainardi, "Detection of atrial fibrillation episodes using a wristband device," *Physiol. Meas.*, vol. 38, no. 5, pp. 787–799, Apr. 2017.
- [62] T. Conroy, J. H. Guzman, B. Hall, G. Tsouri, and J.-P. Couderc, "Detection of atrial fibrillation using an earlobe photoplethysmographic sensor," *Physiol. Meas.*, vol. 38, no. 10, pp. 1906–1918, Sep. 2017.
- [63] E. Gil, P. Laguna, J. Martinez, O. Barquero-Perez, A. Garcia-Alberola, and L. Sörnmo, "Heart rate turbulence analysis based on photoplethysmography," *IEEE Trans. Biomed. Eng.*, vol. 60, pp. 3149–3155, Nov. 2013.
- [64] A. Solosenko, A. Petrenas, and V. Marozas, "Photoplethysmography-based method for automatic detection of premature ventricular contractions," *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 5, pp. 662–669, Oct. 2015.
- [65] L. Polanà, L. Mestha, D. Huang, and J.-P. Couderc, "Method for classifying cardiac arrhythmias using photoplethysmography," in *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2015, pp. 6574–6577.

- [66] M. Lemay, S. Fallet, P. Renevey, C. Leupi, E. Pruvot, and J. M. Vesin, "Wrist-located optical device for atrial fibrillation screening: A clinical study on twenty patients," in *Computing in Cardiology*, Sep. 2016, pp. 681–684.
- [67] S. Fallet, M. Lemay, , P. Renevey, C. Leupi, E. Pruvot, and J. M. Vesin, "An adaptive organization index to characterize atrial fibrillation using wrist-type photoplethysmographic signals," in *Computing in Cardiology*, Sep. 2016, pp. 337–340.
- [68] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, 1994, pp. 359–370.
- [69] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.
- [70] T. H. Everett, J. R. Moorman, L.-C. Kok, J. G. Akar, and D. E. Haines, "Assessment of global atrial fibrillation organization to optimize timing of atrial defibrillation," *Circulation*, vol. 103, no. 23, pp. 2857–2861, Jun. 2001.
- [71] A. Buttu, E. Pruvot, J. Van Zaen, A. Viso, A. Forclaz, P. Pascale, S. M. Narayan, and J.-M. Vesin, "Adaptive frequency tracking of the baseline ECG identifies the site of atrial fibrillation termination by catheter ablation," *Biomed. Signal Process. Control*, vol. 8, no. 6, pp. 969–980, Nov. 2013.
- [72] C. Bandt and B. Pompe, "Permutation entropy: A natural complexity measure for time series," *Phys. Rev. Lett.*, vol. 88, no. 17, p. 174102, Apr. 2002.
- [73] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, "Permutation entropy and its main biomedical and econophysics applications: A review," *Entropy*, vol. 14, pp. 1553–1577, Aug. 2012.
- [74] I. A. Rezek and S. J. Roberts, "Stochastic complexity measures for physiological signal analysis," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 9, pp. 1186–1191, Sep. 1998.
- [75] Electrophysiology, Task Force of the European Society of Cardiology the North American Society of Pacing, "Heart rate variability standards of measurement, physiological interpretation, and clinical use," *Circulation*, vol. 93, no. 5, pp. 1043–1065, Jan. 1996.
- [76] S. T. Lawless, "Crying wolf: false alarms in a pediatric intensive care unit," *Crit. Care Med.*, vol. 22, no. 6, pp. 981–985, Jun. 1994.
- [77] M.-C. Chambrin, "Alarms in the intensive care unit: how can the number of false alarms be reduced?" *Crit. Care*, vol. 5, pp. 184–188, May 2001.
- [78] S. Fallet, S. Yazdani, and J. M. Vesin, "A multimodal approach to reduce false arrhythmia alarms in the intensive care unit," in *Computing in Cardiology*, Sep. 2015, pp. 277–280.
- [79] S. Fallet, S. Yazdani, and J.-M. Vesin, "False arrhythmia alarms reduction in the intensive care unit: a multimodal approach," *Physiol. Meas.*, vol. 37, no. 8, pp. 1217–1232, Aug. 2016.
- [80] A. Mäkitvirta, E. Koski, A. Kari, and T. Sukuvaara, "The median filter as a preprocessor for a patient monitor limit alarm system in intensive care," *Comput. Methods Programs Biomed.*, vol. 34, pp. 139–144, Feb. 1991.

- [81] A. Aboukhalil, L. Nielsen, M. Saeed, R. G. Mark, and G. D. Clifford, "Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform," *J. Biomed. Inform.*, vol. 41, pp. 442–451, Jun. 2008.
- [82] J. Sun, A. Reisner, and R. Mark, "A signal abnormality index for arterial blood pressure waveforms," *Comput. in Cardiol.*, vol. 33, pp. 13–16, 2006.
- [83] W. Zong, G. B. Moody, and R. G. Mark, "Reduction of false arterial blood pressure alarms using signal quality assessment and relationships between the electrocardiogram and arterial blood pressure," *Med. Biol. Eng. Comput.*, vol. 42, pp. 698–706, Sep. 2004.
- [84] Q. Li and G. D. Clifford, "Suppress false arrhythmia alarms of ICU monitors using heart rate estimation based on combined arterial blood pressure and ECG analysis," in *ICBBE*, 2008, pp. 2185–2187.
- [85] —, "Dynamic time warping and machine learning for signal quality assessment of pulsatile signals," *Physiol. Meas.*, vol. 33, pp. 1491–1501, Sep. 2012.
- [86] —, "Signal quality and data fusion for false alarm reduction in the intensive care unit," *J. Electrocardiol.*, vol. 45, pp. 596–603, Nov. 2012.
- [87] J. Behar, J. Oster, Q. Li, and G. Clifford, "ECG signal quality during arrhythmia and its application to false alarm reduction," *IEEE Trans. Biomed. Eng.*, vol. 60, pp. 1660–1666, Jun. 2013.
- [88] R. Salas-Boni, Y. Bai, P. R. E. Harris, B. J. Drew, and X. Hu, "False ventricular tachycardia alarm suppression in the ICU based on the discrete wavelet transform in the ECG signal," *J. Electrocardiol.*, vol. 47, pp. 775–780, 2014.
- [89] Q. Li, C. Rajagopalan, and G. D. Clifford, "Ventricular fibrillation and tachycardia classification using a machine learning approach," *IEEE Trans. Biomed. Eng.*, vol. 61, pp. 1607–1613, 2014.
- [90] G. D. Clifford, I. Silva, B. Moody, Q. Li, D. Kella, A. Shahin, T. Kooistra, D. Perry, and R. G. Mark, "The PhysioNet/Computing in Cardiology Challenge 2015: Reducing false arrhythmia alarms in the ICU," in *Computing in Cardiology*, Sep. 2015, pp. 273–276.
- [91] G. D. Clifford, I. Silva, B. Moody, Q. Li, D. Kella, A. Chahin, Tristan Kooistra, D. Perry, and R. G. Mark, "False alarm reduction in critical care," *Physiol. Meas.*, vol. 37, no. 8, pp. E5–E23, Jul. 2016.
- [92] S. Yazdani and J.-M. Vesin, "Adaptive mathematical morphology for QRS fiducial points detection in the ECG," *Computing in Cardiology*, vol. 41, pp. 725–728, 2014.
- [93] —, "Extraction of QRS fiducial points from the ECG using adaptive mathematical morphology," *Digit. Signal Process.*, vol. 56, no. C, pp. 100–109, Sep. 2016.
- [94] S. Arberet, M. Lemay, P. Renevey, J. Solà, O. Grossenbacher, D. Andries, C. Sartori, and M. Bertschi, "Photoplethysmography-based ambulatory heartbeat monitoring embedded into a dedicated bracelet," in *Computing in Cardiology*, 2013, pp. 935–938.
- [95] S. Fallet, S. Yazdani, and J.-M. Vesin, "A multimodal approach to reduce false arrhythmia alarms in the intensive care unit," in *Computing in Cardiology*, 2015, pp. 277–280.
- [96] W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, "Algorithmic Principles of Remote PPG," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 7, pp. 1479–1491, Jul. 2017.

- [97] Y. Sun, C. Papin, V. Azorin-Peris, R. Kalawsky, S. Greenwald, and S. Hu, "Use of ambient light in remote photoplethysmographic systems: comparison between a high-performance camera and a low-cost webcam," *J. Biomed. Opt.*, vol. 17, no. 3, pp. 0370051–03700510, Mar. 2012.
- [98] D. J. McDuff, E. B. Blackford, and J. R. Estepp, "The impact of video compression on remote cardiac pulse measurement using imaging photoplethysmography," in *2017 12th IEEE International Conference on Automatic Face Gesture Recognition*, May 2017, pp. 63–70.
- [99] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Express*, vol. 18, no. 10, pp. 10762–10774, May 2010.
- [100] —, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7–11, Jan. 2011.
- [101] M. Lewandowska, J. Ruminski, T. Kocejko, and J. Nowak, "Measuring pulse rate with a webcam - A non-contact method for evaluating cardiac activity," in *2011 Federated Conference on Computer Science and Information Systems*, Sep. 2011, pp. 405–410.
- [102] D. McDuff, S. Gontarek, and R. W. Picard, "Improvements in remote cardiopulmonary measurement using a five band digital camera," *Biomed. Opt. Express*, vol. 61, no. 10, pp. 2593–2601, Oct. 2014.
- [103] G. de Haan and V. Jeanne, "Robust pulse rate from chrominance-based rPPG," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2878–2886, Oct. 2013.
- [104] G. d. Haan and A. van Leest, "Improved motion robustness of remote-PPG by using the blood volume pulse signature," *Physiol. Meas.*, vol. 35, no. 9, pp. 1913–1926, Aug. 2014.
- [105] L. Feng, L.-M. Po, X. Xu, Y. Li, and R. Ma, "Motion-resistant remote imaging photoplethysmography based on the optical properties of skin," *EEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 5, pp. 879–891, May 2015.
- [106] W. Wang, S. Stuijk, and G. de Haan, "A novel algorithm for remote photoplethysmography: Spatial subspace rotation," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 9, pp. 1974–1984, Sep. 2016.
- [107] G. R. Tsouri and Z. Li, "On the benefits of alternative color spaces for noncontact heart rate measurements using standard red-green-blue cameras," *J. Biomed. Opt.*, vol. 20, no. 4, pp. 048002–048002, Apr. 2015.
- [108] F. Bousefsaf, C. Maaoui, and A. Pruski, "Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate," *Biomed. Signal Process. Control*, vol. 8, no. 6, pp. 568–574, Nov. 2013.
- [109] L. Wei, Y. Tian, Y. Wang, T. Ebrahimi, and T. Huang, "Automatic webcam-based human heart rate measurements using laplacian eigenmap," in *Computer Vision – ACCV 2012*, ser. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, Nov. 2012, pp. 281–292.
- [110] X. Li, J. Chen, G. Zhao, and M. Pietikäinen, "Remote heart rate measurement from face videos under realistic situations," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2014, pp. 4264–4271.

- [111] H.-Y. Wu, M. Rubinstein, E. Shih, J. Gutttag, F. Durand, and W. Freeman, "Eulerian video magnification for revealing subtle changes in the world," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 65:1–65:8, Jul. 2012.
- [112] P. V. Rouast, M. T. Adam, R. Chiong, D. Cornforth, and E. Lux, "Remote heart rate measurement using low-cost RGB face video: a technical literature review," *Front. Comput. Sci.*, Aug. 2016.
- [113] F. Willemin, "Robust visual tracking using feature selection," Master's thesis, Ecole Polytechnique Fédérale de Lausanne, 2014.
- [114] S. Hare, S. Golodetz, A. Saffari, V. Vineet, M. M. Cheng, S. L. Hicks, and P. H. S. Torr, "Struck: Structured output tracking with kernels," *IEEE Trans. Pattern. Anal. Mach. Intell.*, vol. 38, no. 10, pp. 2096–2109, Oct. 2016.
- [115] B. D. Holton, K. Mannapperuma, P. J. Lesniewski, and J. C. Thomas, "Signal recovery in imaging photoplethysmography," *Physiol. Meas.*, vol. 34, no. 11, pp. 1499–1511, Nov. 2013.
- [116] S. Fallet, V. Moser, F. Braun, and J. M. Vesin, "Imaging photoplethysmography: What are the best locations on the face to estimate heart rate?" in *Computing in Cardiology*, 2016, pp. 341–344.
- [117] Sungjun Kwon, Jeehoon Kim, Dongseok Lee, and Kwangsuk Park, "ROI analysis for remote photoplethysmography on facial video," in *Engineering in Medicine and Biology Society (EMBC), 2015*, 2015, pp. 4938–4941.
- [118] M. A. Hassan, G. S. Malik, N. Saad, B. Karasfi, Y. S. Ali, and D. Fofi, "Optimal source selection for image photoplethysmography," in *2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, May 2016, pp. 1–5.
- [119] S. Kwon, H. Kim, and K. S. Park, "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone," in *Conf Proc IEEE Eng Med Biol Soc*, 2012, pp. 2174–2177.
- [120] K.-Z. Lee, P.-C. Hung, and L.-W. Tsai, "Contact-free heart rate measurement using a camera," in *Ninth Conference on Computer and Robot Vision*, 2012, pp. 147–152.
- [121] W. Wang, S. Stuijk, and G. de Haan, "Exploiting spatial redundancy of image sensor for motion robust rPPG," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 415–425, Feb. 2015.
- [122] B. Han, K. Ivanov, L. Wang, and Y. Yan, "Exploration of the optimal skin-camera distance for facial photoplethysmographic imaging measurement using cameras of different types," in *EAI Endorsed Transactions on Ambient Systems*, no. 11, Dec. 2015, pp. 186–189.
- [123] W. J. Jiang, S. C. Gao, P. Wittek, and L. Zhao, "Real-time quantifying heart beat rate from facial video recording on a smart phone using Kalman filters," in *2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom)*, Oct. 2014, pp. 393–396.
- [124] L. Iozzia, L. Cerina, and L. Mainardi, "Relationships between heart-rate variability and pulse-rate variability obtained from video-PPG signal using ZCA," *Physiol. Meas.*, vol. 37, no. 11, pp. 1934–1944, Nov. 2016.
- [125] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. A. Clifton, and C. Pugh, "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models," *Physiol. Meas.*, vol. 35, no. 5, pp. 807–831, May 2014.

- [126] S. Fallet, L. Mirmohamadsadeghi, V. Moser, F. Braun, and J. M. Vesin, "Real-time approaches for heart rate monitoring using imaging photoplethysmography," in *Computing in Cardiology*, 2016, pp. 345–348.
- [127] S. K. Mitra, *Digital signal processing: a computer-based approach*, ser. McGraw-Hill series in electrical and computer engineering. New York: McGraw-Hill, 1998.
- [128] R. Badeau, G. Richard, and B. David, "Sliding window adaptive SVD algorithms," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 1–10, Jan. 2004.
- [129] M. Clint and A. Jennings, "A simultaneous iteration method for the unsymmetric eigenvalue problem," *J. Inst. Math. Appl.*, vol. 8, no. 1, pp. 111–121, Aug. 1971.
- [130] E. Aboutanios and B. Mulgrew, "Iterative frequency estimation by interpolation on Fourier coefficients," *IEEE Trans. Signal Process.*, vol. 53, pp. 1237–1242, Mar. 2005.
- [131] L. Mirmohamadsadeghi and J. M. Vesin, "Real-time multi-signal frequency tracking with a bank of notch filters to estimate the respiratory rate from the ECG," *Physiol Meas.*, vol. 37, no. 9, pp. 1573–1587, Aug. 2016.
- [132] S. Fallet, Y. Schoenenberger, L. Martin, F. Braun, V. Moser, and J. M. Vesin, "Imaging photoplethysmography: A real-time signal quality index," in *Computing in Cardiology*, 2017.
- [133] J. Sztajzel, "Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system," *Swiss Med. Wkly.*, vol. 134, no. 35-36, pp. 514–522, Sep. 2004.
- [134] A. Schäfer and J. Vagedes, "How accurate is pulse rate variability as an estimate of heart rate variability?" *Int. J. Cardiol.*, vol. 166, no. 1, pp. 15–29, Jun. 2013.
- [135] A. Suhrbier, R. Heringer, T. Walther, H. Malberg, and N. Wessel, "Comparison of three methods for beat-to-beat-interval extraction from continuous blood pressure and electrocardiogram with respect to heart rate variability analysis," *Biomedizinische Technik/Biomedical Engineering*, vol. 51, no. 2, pp. 70–76, Jul. 2006.
- [136] K. Charlot, J. Cornolo, J. V. Brugniaux, J. P. Richalet, and A. Pichon, "Interchangeability between heart rate and photoplethysmography variabilities during sympathetic stimulations," *Physiol. Meas.*, vol. 30, no. 12, pp. 1357–1369, Dec. 2009.
- [137] F. Chang, C. Chang, C. Chiu, S. Hsu, and Y. Lin, "Variations of HRV analysis in different approaches," in *Computers in Cardiology*, Sep. 2007, pp. 17–20.
- [138] E. Gil, M. Orini, R. Bailón, J. M. Vergara, L. Mainardi, and P. Laguna, "Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions," *Physiol. Meas.*, vol. 31, no. 9, pp. 1271–1290, Sep. 2010.
- [139] G. Lu, F. Yang, J. A. Taylor, and J. F. Stein, "A comparison of photoplethysmography and ECG recording to analyse heart rate variability in healthy subjects," *J. Med. Eng. Technol.*, vol. 33, no. 8, pp. 634–641, Nov. 2009.
- [140] P. S. McKinley, P. A. Shapiro, E. Bagiella, M. M. Myers, R. E. De Meersman, I. Grant, and R. P. Sloan, "Deriving heart period variability from blood pressure waveforms," *J. Appl. Physiol.*, vol. 95, no. 4, pp. 1431–1438, Oct. 2003.

- [141] N. Selvaraj, A. Jaryal, J. Santhosh, K. K. Deepak, and S. Anand, "Assessment of heart rate variability derived from finger-tip photoplethysmography as compared to electrocardiography," *J. Med. Eng. Technol.*, vol. 32, no. 6, pp. 479–484, Jan. 2008.
- [142] N. D. Giardino, P. M. Lehrer, and R. Edelberg, "Comparison of finger plethysmograph to ECG in the measurement of heart rate variability," *Psychophysiology*, vol. 39, no. 2, pp. 246–253, Mar. 2002.
- [143] S. Carrasco, R. González, J. Jiménez, R. Román, V. Medina, and J. Azpiroz, "Comparison of the heart rate variability parameters obtained from the electrocardiogram and the blood pressure wave," *J. Med. Eng. Technol.*, vol. 22, no. 5, pp. 195–205, Oct. 1998.
- [144] I. Constant, D. Laude, I. Murat, and J. L. Elghozi, "Pulse rate variability is not a surrogate for heart rate variability," *Clin. Sci.*, vol. 97, no. 4, pp. 391–397, Oct. 1999.
- [145] N. K. Kristiansen, J. Fleischer, M. S. Jensen, K. S. Andersen, and H. Nygaard, "Design and evaluation of a handheld impedance plethysmograph for measuring heart rate variability," *Med. Biol. Eng. Comput.*, vol. 43, no. 4, pp. 516–521, Jul. 2005.
- [146] S. Lu, H. Zhao, K. Ju, K. Shin, M. Lee, K. Shelley, and K. H. Chon, "Can photoplethysmography variability serve as an alternative approach to obtain heart rate variability information?" *J. Clin. Monit. Comput.*, vol. 22, no. 1, pp. 23–29, Jan. 2008.
- [147] S. Yazdani, S. Fallet, and J. M. Vesin, "A novel short-term event extraction algorithm for biomedical signals," *IEEE Trans. Biomed. Eng.*, 2017.
- [148] S. Yazdani and J. M. Vesin, "A novel preprocessing tool to enhance ECG R-wave extraction," in *Computing in Cardiology*, 2016, pp. 633–636.
- [149] M. J. Drinnan, J. Allen, and A. Murray, "Relation between heart rate and pulse transit time during paced respiration," *Physiol. Meas.*, vol. 22, no. 3, pp. 425–432, Aug. 2001.
- [150] S. Dawson, R. Panerai, and J. Potter, "Should one use electrocardiographic or Finapres-derived pulse intervals for calculation of cardiac baroreceptor sensitivity?" *Blood Press. Monit.*, vol. 3, no. 5, pp. 315–320, Oct. 1998.
- [151] J.-P. Couderc, S. Kyal, L. K. Mestha, B. Xu, D. R. Peterson, X. Xia, and B. Hall, "Detection of atrial fibrillation using contactless facial video monitoring," *Heart Rhythm*, vol. 12, no. 1, pp. 195–201, Jan. 2015.
- [152] W. Karlen, A. Garde, D. Myers, C. Scheffer, J. Ansermino, and G. Dumont, "Estimation of respiratory rate from photoplethysmographic imaging videos compared to pulse oximetry," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 4, pp. 1331–1338, Jul. 2015.
- [153] J. Perk, G. De Backer, H. Gohlke, I. Graham, Ž. Reiner, M. Verschuren, C. Albus, P. Benlian, G. Boysen, R. Cifkova *et al.*, "European guidelines on cardiovascular disease prevention in clinical practice (version 2012)," *Eur. Heart J.*, vol. 33, no. 13, pp. 1635–1701, Sep. 2012.
- [154] L. Mirmohamadsadeghi and J. M. Vesin, "Estimating the real-time respiratory rate from the ECG with a bank of notch filters," in *Computing in Cardiology*, 2015, pp. 581–584.
- [155] P. H. Charlton, T. Bonnici, L. Tarassenko, D. A. Clifton, R. Beale, and P. J. Watkinson, "An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram," *Physiol Meas*, vol. 37, no. 4, pp. 610–626, Apr. 2016.

- [156] L. Mirmohamadsadeghi and J.-M. Vesin, "Respiratory rate estimation from the ECG using an instantaneous frequency tracking algorithm," *Biomed. Signal Process. Control*, vol. 14, pp. 66–72, Nov. 2014.
- [157] L. Mirmohamadsadeghi, S. Fallet, V. Moser, F. Braun, and J. M. Vesin, "Real-time respiratory rate estimation using imaging photoplethysmography inter-beat intervals," in *Computing in Cardiology*, 2016, pp. 861–864.
- [158] R. Bailón, L. Sörnmo, and P. Laguna, *Advanced Methods and Tools for ECG Data Analysis*. Artech House, 2006, ch. 8: ECG- Derived Respiratory Frequency Estimation, pp. 215–243.
- [159] R. L. Horner, D. Brooks, L. F. Kozar, K. Gan, and E. A. Phillipson, "Respiratory-related heart rate variability persists during central apnea in dogs: mechanisms and implications," *J. Appl. Physiol.*, vol. 78, no. 6, pp. 2003–2013, June 1995.
- [160] A. Buttu, "Novel ECG and intracardiac electrograms signal processing schemes for predicting the outcome of atrial fibrillation catheter ablation," Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne, 2014.
- [161] E. Sazonov and M. R. Neuman, *Wearable sensors: fundamentals, implementation and applications*. Elsevier, Aug. 2014.
- [162] F. Zhao, M. Li, Z. Jiang, J. Z. Tsien, and Z. Lu, "Camera-based, non-contact, vital-signs monitoring technology may provide a way for the early prevention of SIDS in infants," *Front. Neurol.*, vol. 7, p. 236, Dec. 2016.
- [163] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.
- [164] E. Alpaydin, *Introduction to machine learning*. MIT press, 2014.
- [165] J. Surowiecki, "The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business," *Economies, Societies and Nations*, vol. 296, 2004.
- [166] —, *The wisdom of crowds*. Anchor, 2005.
- [167] B. S, "Gesundheit der neugeborenen 2008. frühgeburten, mehrlingsschwangerschaften und wachstumsretardierung." Bern: Bundesamt für Statistik BFS Aktuell, 2010.
- [168] M. Adams and H. U. Bucher, "Néonatalogie: Un début précoce dans la vie: qu'apporte un registre national?" *Forum Med Suisse*, vol. 13, no. 3, pp. 35–37, 2013.
- [169] M. Adams, C. Borradori-Tolsa, M. Bickle-Graz, S. Grunt, P. Weber, A. Capone Mori, F. Bauder, C. Hagmann, G. Natalucci, R. Pfister, and B. Latal, "Follow-up assessment of high-risk newborns in Switzerland," *Paediatrica*, vol. 25, no. 5, pp. 8–10, 2014.
- [170] L. A. M. Aarts, V. Jeanne, J. P. Cleary, C. Lieber, J. S. Nelson, S. Bambang Oetomo, and W. Verkruysse, "Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit - A pilot study," *Early Hum. Dev.*, vol. 89, no. 12, pp. 943–948, Dec. 2013.
- [171] M. Villarroel, A. Guazzi, J. Jorge, S. Davis, P. Watkinson, G. Green, A. Shenvi, K. McCormick, and L. Tarassenko, "Continuous non-contact vital sign monitoring in neonatal intensive care unit," *Healthc. Technol. Lett.*, vol. 1, no. 3, pp. 87–91, Sep. 2014.
- [172] M. K. Park, *Park's pediatric cardiology for practitioners*, 6th ed., ser. Pediatrics. Philadelphia, PA: Elsevier Saunders, 2014.

-
- [173] F. Braun, A. Lemkaddem, V. Moser, S. Dasen, O. Grossenbacher, and M. Bertschi, “Contactless respiration monitoring in real-time via a video camera,” in *EMBECE & NBC 2017*, ser. IFMBE Proceedings, Jun. 2017, pp. 567–570.

Sibylle Fallet

Address: Chemin de Bel-Air 1, 2000 Neuchâtel

Date of birth: 26.06.1990

Phone: +41795847658

Nationality: Swiss

Email: sibyllefallet@gmail.com

Education

Ph.D. student in Biomedical Signal Processing Applied Signal Processing Group, Swiss Federal Institute of Technology, Lausanne. Under the direction of Dr. Jean-Marc Vesin, EPFL	2013-present
Master of Science (M. Sc) in Bioengineering Swiss Federal Institute of Technology, Lausanne. Minor in Biomedical Technologies.	2011-2013
Bachelor of Science (B. Sc) in Life Sciences and Technology Swiss Federal Institute of Technology, Lausanne.	2008-2011
High school degree Lycée Blaise-Cendrars (La Chaux-de-Fonds, NE).	2005-2008

Research experience

Swiss Federal Institute of technology, Lausanne, Switzerland <i>PhD researcher</i> <ul style="list-style-type: none">• Conduct extensive data analysis using Matlab.• Develop signal processing tools for innovative cardiovascular monitoring applications:<ul style="list-style-type: none">◦ Robust heart rate estimation using photoplethysmography.◦ Video-based vital signs monitoring.◦ Arrhythmias detection using photoplethysmography.• Present results in conferences.• Write articles.• Collaborate actively with companies and hospitals.	2013-present
--	--------------

Professional experience

Motilis Medica S.A., Neuchâtel, Switzerland <i>Master's Researcher under the supervision of Vincent Schlageter (Motilis Medica S.A.) and Jean-Marc Vesin (EPFL)</i> <ul style="list-style-type: none">• Processed the trajectories data from ingestible telemetric capsules to monitor gastrointestinal motility.• Involved in the analysis of data from a clinical study to correlate motility patterns and sleep stages.• Collaborated with M.D.s from Aarhus University Hospital, Denmark.	2012-2013
Help Desk DIT - Call center, Swiss Federal Institute of technology, Lausanne, Switzerland <i>Student assistant</i> <ul style="list-style-type: none">• Provided support to end-users with short response time.• Acquired stress management and communication skills.	2012-2013 (half-day/week)

Voluntary work -----

Health Promotion and System Strengthening Project, Tanzania

2013 (sept-nov)

Volunteer EPFL biomedical engineer

- Conducted an assessment of medical waste management.
- Audited 70 Tanzanian health care facilities.
- Collaborated with Swiss Tropical Health Institute and EssentialMed.
- Wrote a technical report with recommendations, which was given to the Tanzanian Ministry of Health to support their decision-making process.

Personal skills and competences -----

Languages: French: mother tongue, English: advanced, German: basic.

Computer and programming skills: Matlab, C++ , LaTeX.

Awards -----

- 2016: **Best poster award** for the poster entitled: *Imaging photoplethysmography: what are the best locations on the face to estimate heart rate?*, Computing in Cardiology 2016, 11-14 September, Vancouver, Canada.
- 2015: **First place PhysioNet/Computing in Cardiology Challenge 2015: Reducing False Arrhythmia Alarms in the ICU.** The 2015 PhysioNet/CinC Challenge aimed to encourage the development of algorithms to reduce the incidence of false alarms in the Intensive Care Unit (ICU).
- 2013: **Best poster award** for the poster entitled: *Post-processing of the trajectories of telemetric capsules for gastrointestinal motility monitoring*, EPFL Master in Bioengineering.

Publications

Journal Papers

S. Fallet, M. Lemay, P. Renevey, C. Leupi, E. Pruvot and J.-M. Vesin, “Can one detect atrial fibrillation using a wrist-type photoplethysmographic device?,” *submitted to Medical & Biological Engineering & Computing*,

S. Yazdani, S. Fallet and J.-M. Vesin, “A Novel Short-term Event Extraction Algorithm for Biomedical Signals,” *IEEE Transactions on Biomedical Engineering*, 2017

S. Fallet and J.-M. Vesin, “Robust heart rate estimation using wrist-type photoplethysmographic signals during physical exercise: An approach based on adaptive filtering,” *Physiological Measurement*, vol. 38 no 2, pp. 155–170, January 2017

S. Fallet, S. Yazdani, and J.-M. Vesin, “False arrhythmia alarms reduction in the intensive care unit: A multimodal approach,” *Physiological Measurement*, vol. 37 no 8, pp. 1217–1232, July 2016

J.J. Saugy, L. Schmitt, S. Fallet, R. Faiss, J.-M. Vesin, M. Bertschi, R. Heinzer and G.P. Millet, “Sleep disordered breathing during live high-train low in normobaric versus hypobaric hypoxia,” *High Altitude Medicine & Biology*, vol. 17 no 3, pp. 233–238, July 2016

Conference Papers

S. Fallet, Y. Schoenenberger, L. Martin, F. Braun, V. Moser and J.-M. Vesin, “Imaging photoplethysmography: A real-time signal quality index,” *2017 Computing in Cardiology Conference (CinC)*, Rennes, France, September 2017.

S. Fallet, M. Lemay, P. Renevey, C. Leupi, E. Pruvot and J.-M. Vesin, “An adaptive organization index to characterize atrial fibrillation using wrist-type photoplethysmographic signals,” *2016 Computing in Cardiology Conference (CinC)*, Vancouver, Canada, September 2016.

M. Lemay, S. Fallet, P. Renevey, C. Leupi, E. Pruvot and J.-M. Vesin, “Wrist-located optical device for atrial fibrillation screening: A clinical study on twenty patients,” *2016 Computing in Cardiology Conference (CinC)*, Vancouver, Canada, September 2016.

S. Fallet, L. Mirmohamadsadeghi, M. Moser, F. Braun and J.-M. Vesin, “Real-time approaches for heart rate monitoring using imaging photoplethysmography,” *2016 Computing in Cardiology Conference (CinC)*, Vancouver, Canada, September 2016.

S. Fallet, M. Moser, F. Braun and J.-M. Vesin, “Imaging photoplethysmography: What are the best locations on the face to estimate heart rate?,” *2016 Computing in Cardiology Conference (CinC)*, Vancouver, Canada, September 2016.

L. Mirmohamadsadeghi, S. Fallet, M. Moser, F. Braun and J.-M. Vesin, “Real-time respiratory rate estimation using imaging photoplethysmography inter-beat intervals,” *2016 Computing in Cardiology Conference (CinC)*, Vancouver, Canada, September 2016.

S. Fallet, S. Yazdani, and J.-M. Vesin, “A multimodal approach to reduce false arrhythmia alarms in the intensive care unit,” *2015 Computing in Cardiology Conference (CinC)*, Nice, France, September 2015.

S. Fallet and J.-M. Vesin, “Adaptive frequency tracking for robust heart rate estimation using wrist-type photoplethysmographic signals during physical exercise,” *2015 Computing in Cardiology Conference (CinC)*, Nice, France, September 2015.

L. Mirmohamadsadeghi, S. Fallet, A. Buttu, J. Saugy, T. Rupp, R. Heinzer, J.-M. Vesin and G.P. Millet, “Sleep apnea detection using features from the respiration and the ecg recorded with smart-shirts,” *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Lausanne, Switzerland, October 2014.

Abstracts

S. Fallet, M. Moser, D. Ostojic, T. Karen, F. Braun and J.-M. Vesin, “Video-based heart rate monitoring in neonatal intensive care unit: Preliminary results,” *2017 Biomedical Photonics Network (BMPN)*, Berne, Switzerland, November 2017.

S. Fallet, M. Lemay, P. Renevey, C. Leupi, E. Pruvot and J.-M. Vesin, “Can one detect atrial fibrillation using a wrist-type photoplethysmographic device?,” *2017 joint conference of the European Medical and Biological Engineering Conference (EMBEC) and the Nordic-Baltic Conference on Biomedical Engineering and Medical Physics (NBC)*, Tampere, Finland, June 2017.

